TRACKING INDUCED SEISMICITY IN THE FORT WORTH BASIN, TEXAS AND NORTHERN OKLAHOMA USING LOCAL AND LARGE-N STYLE ARRAYS

Approved by:

_____
Dr. Heather R. DeShon
Professor of Geophysics

_____
Dr. Matthew Hornbach
Professor of Geophysics

_____
Dr. Maria Beatrice Magnani
Professor of Geophysics

_____
Dr. Brian W. Stump
Albritton Professor of Geophysics

_____
Dr. Peter H. Hennings
Research Scientist at University of Texas

TRACKING INDUCED SEISMICITY IN THE FORT WORTH BASIN, TEXAS AND NORTHERN OKLAHOMA USING LOCAL AND LARGE-N STYLE ARRAYS

A Dissertation Presented to the Graduate Faculty of

Dedman College

Southern Methodist University

in

Partial Fulfillment of the Requirements

for the degree of

Doctor of Philosophy

with a

Major in Geophysics

by

Louis Quinones

B.S., Geophysics, Texas A&M University, College Station, TX
Master of Science, Southern Methodist University, TX

December 18, 2021

Copyright (2021)

Louis Quinones

All Rights Reserved

ACKNOWLEDGMENTS

First, I thank my advisor Heather DeShon, along with the members of my committee: Beatrice Magnani, Brian Stump, Matt Hornbach, and Peter Hennings as well as Chris Hayward and Stephen Arrowsmith for all their advice and suggestions that improved the quality of the work as a whole. I also thank my fellow graduate students SeongJu Jeong, Kevin Kwong, Vanshan Wright, and Monique Holt for all of their advice and support without which I wouldn't have been able to complete this dissertation. I thank the members of the TexNet program for both funding this project and for providing access to their seismic instruments and data in the Fort Worth Basin. Similarly, I thank SMU, IRIS PASSCAL, and the USGS for providing the seismic instruments used to monitor the seismicity in the Fort Worth Basin. Much of the methodology work in this thesis has been previously published by me in the publication denoted as Quinones et al. (2019) with co-authorship given to: Heather DeShon, SeongJu Jeong, Paul Ogwari, Oner Sufri, Monique Holt, and Kevin Kwong, and has been rewritten here as the work contained in the publication was completed as part of the research conducted for this dissertation. The work completed in Chapter 4 was adapted from a submitted manuscript with co-authorship given to: Heather DeShon, Theresa Sawi, Stephen Arrowsmith, and Katie Keranen. Any tables or figures shared between the publication and this dissertation have been copied with full permission by all authors.

Quinones, Louis A. B.S., Geophysics, Texas A&M University
M.S., Geophysics, Southern Methodist University

Tracking Induced Seismicity in the Fort Worth Basin, Texas and Northern Oklahoma Using Local and Large-N Style Arrays

Advisor: Professor Heather R. DeShon
Doctor of Philosophy conferred December 18, 2021
Dissertation completed September 1, 2021

Since 2008, earthquake sequences within the Fort Worth Basin (FWB), north Texas, have been linked to wastewater disposal activities related to unconventional shale gas production. Here, I present my work analyzing and cataloging a complete record of the seismicity occurring within the basin over a period of more than a decade (2008-2020). Analysis of the catalog reveals that the earthquakes generally occur within the Precambrian basement along steeply dipping normal faults, and while overall seismicity rates have decreased since 2016, new faults have become active. I observe strong spatial and temporal correlations between the earthquake locations and wastewater disposal well locations and injection volumes, implying that fluid injection activities may be the main driving force of seismicity in the basin.

To analyze the basin-wide injection associated stress changes, I created a coupled geomechanical model of the Fort Worth Basin incorporating geology and fault data shared by our collaborators at TexNet and the Bureau of Economic Geology. Models exploring endmember geologic and tectonic scenarios were created using the COMSOL modeling software and are fully coupled, meaning that both injection associated pore fluid pressure and poroelastic stress changes are calculated. The preferred model shows that while large injection associated stress changes are concentrated in those portions of the basin where both injection volumes and rates

are highest, there are observable far-field stress changes capable of reaching faults at far distances (>15 km) away from injection wells. Further work focusing on the Dallas-Irving sequence, located >15 km from low-to-medium volume injection wells, involved the creation of a high-resolution relative location catalog and template matching work to respectively better estimate the spatiotemporal characteristics and start date of the sequence. The results indicate that the Dallas-Irving sequence likely began in very late 2013 or early 2014 and that the seismicity associated with the sequence is occurring along a NE-SW trending normal fault. Importantly, the Dallas-Irving sequence only became active following the arrival of injection associated stress changes at the sequence site.

Lastly, the dissertation shifts focus to explore newer earthquake detection techniques to understand their usefulness in monitoring and creating earthquake catalogs efficiently and rapidly. The work focuses on the Community Wavefields Demonstration Experiment (CWDE) data set in northern Oklahoma. The local similarity event detection method utilizes the large-N style nodal array to detect events without the need for *a priori* information related to the geology of the study region. The study focuses on the creation of a benchmark earthquake catalog made using traditional analyst-based analysis of the broadband seismic data collected as part of the community experiment. This benchmark catalog is used for comparison with the local similarity nodal detection list. Ultimately, the local similarity technique was able to detect nearly all (97.0%) events recorded within the benchmark broadband catalog and was also able to detect over 1000 events which were not included in the benchmark catalog. Through the analysis of the local similarity results I was able to extract network-wide first P- and S-wave arrival onset time estimates which were used to discern local and regional distance events from the array.

The NTXES catalog is the most complete record of seismicity occurring in the FWB and represents a great starting point for possible future analysis of long-spanning induced seismicity. Continual monitoring of seismicity in the FWB yields insight into both the lifespan of induced seismicity sequences, and also tell us about how seismic sequences may end. Additionally, the evolution of the FWB seismicity could inform us about the effectiveness of injection mitigation techniques considering that, due to a variety of factors, injection activities have gradually been decreasing since their peak in 2012, and likewise seismicity rates have continuously decreased following their peak in 2015. For monitoring purposes, the created event detection list of the CWDE data set serves as a great starting point, but further analysis steps need to be undertaken to convert these results into a full earthquake catalog. Further analysis of the CWDE data set may be extended to other event detection and phase picking techniques and algorithms which could either seek to take advantage of the benchmark catalog for approaches such as template matching or machine learning or instead focus entirely on the nodal array data and benchmark their results on the local similarity results instead. Ultimately, the CWDE data set remains an excellent data set for the purposes of testing a variety of event detection and phase picking techniques which could be incorporated into the workflow analysis of any future network deployment involving the installation of large-N style broadband and nodal station arrays.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

The study of induced seismicity seeks to understand the physical mechanisms through which anthropogenic activities lead to stress changes capable of triggering slip on pre-existing faults. Examples of anthropogenic activities inducing earthquakes can be found all throughout the 20th and 21st century and causal mechanisms range from the extraction of oil and fluids (e.g., Pratt and Johnson, 1926; Segall, 1989; Taira et al., 2018) to the disposal of fluids (e.g., Bardwell, 1966; Ellsworth, 2013; Keranen et al., 2013) to the construction of reservoirs (e.g., Simpson et al., 1988; Talwani, 1997). Since 2008 there has been a significant increase in the overall rate of seismicity throughout the central US across numerous states including Texas (e.g., Frohlich et al., 2016), Oklahoma (e.g., Keranen et al., 2014), Arkansas (e.g., Horton, 2012) and Kansas (e.g., Rubinstein et al., 2018) that is believed to be induced (Ellsworth, 2013). Most earthquakes in these states have been associated with the increase in fluid disposal activities related to ongoing natural gas and oil production procedures, though some events have been linked directly to hydrofracking in tight shale formations (Rubenstein and Mahani, 2015). This dissertation focuses on induced earthquakes in northern Texas and northern Oklahoma associated with wastewater injection, hereafter referred to as fluid injection induced seismicity.

Fluid injection induced seismicity was primarily thought of as occurring along faults which are near (typically <15 km away) fluid disposal wells (Weingarten et al., 2015). The reason that injection activities are considered a causal mechanism of induced seismicity is due to the direct relationship wherein increases in pore fluid pressure led to decreases in normal stress

values (Hubbert and Rubey, 1959). Studies of fluid injection induced seismicity have observed that the earthquakes associated with these sequences are concentrated within the crystalline basement which typically directly underlies the primary fluid disposal formation within that region of seismicity (e.g., Keranen and Weingarten, 2018). These faults are now thought to be critically stressed faults within the active regional stress field which required only small, on the order of 1 MPa or fewer, stress increases to begin slipping and producing seismicity (Zoback and Hickman, 1982; Liu and Zoback, 1997; Walsh and Zoback, 2016).

The two regions of focused study, the Fort Worth Basin in north-central Texas and the Community Wavefields Demonstration Experiment study site located in northern Oklahoma, contain many of the key characteristics associated with fluid injection induced seismicity but differ greatly from one another in terms of how long each region has been monitored, how the networks monitoring each region were designed and installed, and what types of analysis can be achieved using each data set (Figure 1.1). I utilize these two data sets to address questions about fluid injection induced seismicity such as: how do changing long-term trends in fluid injection activities within a basin affect seismicity rates; what is the expected duration of fluid injection induced seismicity on individual faults or within a basin; what are the best network monitoring strategies to study induced seismicity; what role do far-field fluid injection associated stress changes (>15 km) play in the triggering of induced seismicity; and how can more advanced detection and location techniques help us better focus our monitoring efforts and improve our understanding of induced seismicity sequences.

Numerous studies of the seismicity occurring within the Fort Worth Basin (FWB) have linked the onset of the post-2008 earthquake sequences with the increase in fluid injection activities related to the extraction of shale gas in the basin (Frohlich et al., 2010, 2011, 2016;

Frohlich, 2012; Justinic et al., 2013; Hornbach et al., 2015, 2016; Lund Snee and Zoback, 2016; Magnani et al., 2017; Scales et al., 2017; Ogwari et al., 2018; Quinones et al., 2019). Production and subsequent fluid injection activities within the FWB began in the early 21st century as part of the shale gas boom in which natural gas was extracted primarily from the Barnett Shale formation. These production activities were concentrated in the Newark East Field located within the northeastern corner of the FWB. The waste fluids created as part of this natural gas extraction process are injected into the Ellenburger limestone-dolomite formation at injection well sites across the northeast portion of the basin (Bowker, 2007). The production and fluid injection activities in the FWB peaked around 2012, and since then both activities have seen a sharp decrease due to primarily economic factors associated with the extraction of natural gas from the Barnett Shale which was always considered an unconventional natural gas play (Hornbach et al., 2016). Seismicity has been also concentrated in the northeast portion of the basin, as documented in the national earthquake catalog (Figure 1.1, USGS ComCat, last accessed June 2021).

The FWB covers most of the north-central portion of Texas and is defined as a foreland basin created as part of the Ouachita orogeny during the late Paleozoic (Pollastro et al., 2007). The boundaries of the FWB are defined by the Ouachita Thrust Front to the east, by the Muenster and Red River arches to the north, by the Llano uplift to the south, and by a general shallowing of the major geologic formations in the basin to the west as they approach the Bend Arch (e.g., Smye et al. 2021). The major geologic formations of interest for the study of induced seismicity are those associated with shale gas, injection, and seismicity: the Barnett shale is the major shale gas producing formation; the Ellenburger limestone-dolomite formation is the primary fluid disposal unit; and the crystalline basement hosts most of the seismicity. The Precambrian crystalline basement is the lowest structural unit of the FWB and is overlain in

places by the thin Upper Cambrian Riley formation, which is primarily composed of the highly permeable Hickory Sandstone member (Pollastro et al., 2007; Smye et al., 2021). The Hickory Sandstone formation is in turn overlain by the Ordovician age carbonate units that comprise the Ellenburger group. Overlying these are a set of Mississippian and Pennsylvanian age formations that include the Viola and Barnett Shale formation. Above the Barnett lie multiple groups of Pennsylvanian age foreland basin fill units and, unconformably, Cretaceous age coastal plain sediments. A major shift in the geologic column occurs along the eastern boundary of the basin, where the Ouachita orogeny thrusts meta-sedimentary units over the Ordovician carbonates and Mississippian shales (Magnani et al., 2017; Hennings et al., 2019). Recent fault maps indicate that the orientations of most faults in the basin align with these major tectonic boundaries (Magnani et al., 2017; Hennings et al., 2019; Horne et al., 2020). This can be clearly observed along the eastern boundary of the FWB where the majority of fault structures trend in the NNE-SSW direction, paralleling the Ouachita Thrust Front.

Chapter 2 of this dissertation focuses on the earthquake catalogs associated with the FWB and outlines the effort to combine them into one catalog to best describe the seismic history to understand causal factors triggering seismicity. The resulting catalog, hereafter referred to as the North Texas Earthquake Study (NTXES) catalog, represents the most complete record of seismicity, including the highest quality earthquake locations and magnitudes, for the FWB from 2008-2020. The creation of the NTXES catalog involves manual review of all SMU collected waveform data over this period (Justinic et al. 2013; DeShon et al., 2018) along with the creation of new sequence specific velocity models and a regional seismic attenuation curve function for the calculation of local magnitude. The analysis of the NTXES catalog focuses on the following hypotheses. First, because the seismicity occurring in the FWB has a strong spatiotemporal

relationship with the fluid injection volumes in the basin, then the seismicity occurring in the FWB is induced (as posited by Hornbach et al., 2016). To test this hypothesis, I examine the temporal histories of all the seismic sequences within the NTXES catalog and compare them to the local and basin-wide history of fluid injection activities occurring in the FWB. Second, if the seismicity occurring in the FWB shares the previously stated key characteristics of fluid injection induced seismicity, those being that seismicity is occurring near (<15 km) fluid injection wells on critically stressed faults primarily within the crystalline basement, then fluid injection activities and their associated stress changes are the main driving stress changes inducing seismicity in the basin. I test this hypothesis by examining the location and depth distributions recorded for the various sequences, considering their individual earthquake location uncertainties, and comparing those earthquake location distributions to the locations of the fluid injection wells in the basin and the local geology of each seismic sequence. Another aspect of the location and depth analysis of the NTXES catalog comes from the comparison of those sequences that were monitored by a local network of stations of mixed sensor types and those sequences which were only monitored using regional distance (>15 km) stations from across the basin.

The final hypothesis addressed states that if fluid injection activities continue over long enough time periods within a hydraulically permeable geologic formation, then one should expect to observe far-field (>15 km) injection associated pore fluid pressure and poroelastic stress changes. This hypothesis derives from the observation that the most northeastern seismic sequences in the basin, namely the Dallas-Irving and Lake Lewisville sequences analyzed in Chapter 2, do not have nearby fluid injection wells. Given the previous hypothesis, this observation leads to questions about how these sequences could be fluid injection induced

seismicity. Chapter 3 focuses on this hypothesis through focused, multi-disciplinary study of the Dallas-Irving earthquake sequence. In the FWB this is the sequence with the highest associated risk because it occurs in the middle of a major metropolitan center, the cities of Irving and Dallas. The sequence shares many of the characteristics of a fluid injection induced seismicity sequence, but it lacks an associated nearby fluid injection well to provide the assumed direct pore fluid pressure increase that induced slip on the fault (as shown in Ogwari et al., 2018). As such, the work in this chapter focuses on the following hypothesis. If a sequence occurring within a region of known fluid injection induced seismicity lacks a nearby source of injection associated stress changes, then far-field injection associated stress changes, such as poroelasticity, are the main driving stress changes inducing seismicity at the sequence site. Therefore, if a fault only became seismically active following the arrival of injection associated stress changes, then that fault can be considered an example of fluid injection induced seismicity. The first step towards testing these hypotheses was to ensure that I had the highest resolution earthquake locations associated with the sequence. I use the double-difference relocation technique (*tomoDD*, Zhang and Thurber, 2003) to calculate a revised catalog of the Dallas-Irving sequence and better constrain the spatial distribution of the earthquakes and the fault geometry imaged by those earthquakes. Because monitoring began in conjunction with the largest felt earthquakes in the sequence, I utilize a template-matching approach (*EQcorrscan*, Chamberlain et al., 2018) to examine the pre-local monitoring period of seismicity for the Dallas-Irving fault and estimate an onset time for the sequence.

Chapter 3 also includes a coupled geomechanical model of the seismically active portion of the FWB within the COMSOL modeling software to better assess the basin wide injection associated stress changes. The geomechanical model in this case is referred to as a coupled

model as it solved for both the pore fluid pressure and poroelastic stress changes induced by the injection of fluids into the Ellenburger formation. This geomechanical modeling project related to the previous hypothesis as it allowed me to gain an estimate of the far-field injection associated stress changes across the basin, specifically at the Dallas-Irving sequence site, and the time needed for those stress changes to have propagated to far enough distance away from the injection wells to have reached the farthest northeast portion of the basin. The analysis of the coupled geomechanical model focuses on the temporal changes in the injection associated stress at various locations across the basin, but also delves into the parameter testing required to properly assess the modeling results with regards to the hydraulic properties of the major formations in the basin. Using that temporal estimate from the modeling of when the stress changes would have reached the Dallas-Irving fault and the NTXES catalog allows me to gain an initial estimate of whether the stress changes reached the fault prior to the onset of seismicity. Ultimately, with the template matched detection list and the coupled geomechanical modeling results I can better assess my second hypothesis and more definitively determine whether the Dallas-Irving fault shows the full suite of key characteristics of fluid injection induced seismicity.

The FWB data set allows for effective analysis of fluid injection induced seismicity during an ongoing sequence, but its complex monitoring history in a noisy metropolitan area yields a limited data set for analyses requiring a complete history of small magnitude (<1.5) earthquakes. This led to the question - could we have done better, and if so, what could have been gained in the process. To answer these questions, I use the Community Wavefields Demonstration Experiment (CWDE) to explore modern event detection techniques on state-of-the-art nodal arrays for earthquake sequences believed to be fluid-injection induced. The CWDE

was an experiment led by the Incorporated Research Institutions for Seismology (IRIS) in the summer of 2016 whose focus was to test the usefulness of recording the full seismic wavefield (Sweet et al., 2018). The experiment consisted of a deployment of 18 broadband Guralp CMG-3T sensors, 9 infrasound Hyperion IFS3311 microbarometer stations, and 363 three-component Fairfield Nodal ZLand nodal sensors. The experiment site, a 5 x 13 km area in northern Oklahoma, was selected because of active seismicity allowing for the detection of many earthquakes at close distances to the array. The temporal history of the array can be separated into two periods: the first approximately 35 days of the experiment when all stations were active (06/21/16 – 07/27/16), and a remaining approximately five-month period when only the broadband and infrasound instruments were active (06/18/16 – 11/13/16). The 18 broadband sensors, alongside 9 co-located infrasound sensors, were deployed using the novel Golay array distribution which consisted of six sets of mini arrays each containing three stations. The purpose of this array distribution was to increase the low frequency (<5 Hz) recording capabilities of the array to better observe the full seismic wavefield. The station distribution of the nodal sensors was also unique in that it incorporated both linear deployments of stations along roads in the study area and the construction of a gradiometer array. The gradiometer consisted of approximately 112 nodal stations deployed in a seven-layer rectangular pattern over an 800 x 800 m area with each layer of the gradiometer consisting of 16 nodes. Like the Golay array configuration of the broadband sensors, the purpose of the gradiometer configuration was chosen to better measure the seismic wavefield. The remaining 251 nodal stations were deployed in linear configurations along three roads, one in the east-west direction spanning ~13 km and two in the north-south direction each spanning ~4.8 km in length which were ~3 km apart from one another. The nodal sensors along these three seismic lines were spaced at 100 m intervals with

the exception being that at the intersection points of the seismic lines where the station spacing lowered to as little as 33 m.

The study region of northern Oklahoma is part of the northern Anadarko shelf and lies almost directly west of the Nemaha Uplift (Johnson, 2008). The oldest formations in this region are the Precambrian crystalline basement rocks whose top lies at ~3 km (~10,000 ft). Like in the FWB, the seismically active faults in this region are also basement-rooted faults, which are hydraulically connected to the overlying rock units (Schoenball and Ellsworth, 2017; Qin et al., 2019). However, unlike the FWB, the seismically active faults appear to be closely spaced to one another and the overall seismicity rates throughout Oklahoma far surpass those recorded in north Texas (Schoenball and Ellsworth, 2017). Overlying the Precambrian basement rocks is a set of Ordovician age carbonate units which are dominated by the Arbuckle limestone-dolomite group, genetically identical to the Ellenburger formation in Texas; like the Ellenburger formation, this unit is the main fluid disposal formation for northern Oklahoma. Above the Arbuckle lies a mixture of Mississippian, Devonian, and Silurian age sedimentary rocks that include limestone and shale units with the Woodford Shale formation acting as the main production unit for enhanced oil recovery in this part of Oklahoma (Cardott, 2017). Pennsylvanian age shales associated with the growth of the Nemaha Uplift during the Ouachita orogeny and other mountain building events occurring during this period overlie these units. Lastly are the Permian formations of primarily shale, limestone, and some sandstone.

The study region surrounding the CWDE array has recorded seismicity with compelling characteristics of fluid injection induced seismicity and well-located earthquake catalogs exist for the region (Schoenball and Ellsworth, 2017). The question I sought to explore using the CWDE array was given a modern deployment of a large-N style array, do recently developed

event detection techniques provide a significant benefit to the earthquake monitoring process? In Chapter 4, I compare traditional analyst-based earthquake catalog to an automated approach that uses the large amounts of waveform data associated with large-N style arrays. I also explore the ability of this automated detection technique to create a complete catalog of seismicity for the study area without manual review. This work involves creating both the traditional analyst-based earthquake catalog and the automated event detection list due to neither type of catalog nor detection list previously existing for the CWDE array data set. The analyst-based earthquake catalog is created using the broadband sensor data and follows the same workflow described in the creation of the NTXES catalog in the FWB (described in Chapter 2). The automated event detection list is created using the local similarity technique, which is designed to use large volumes of waveform data collected by densely spaced stations (Li et al., 2018). Following the creation of both the analyst-based catalog and the local similarity detection list, I compare the recording capabilities of both approaches in terms of the manual and computational hours needed to use each technique and the ability of each technique to create a complete catalog of seismicity for the local area around the CWDE array. I then discuss the possible steps that can be taken to convert the local similarity results from a list of event detection times to a full earthquake catalog complete with estimates of earthquake locations and magnitudes.

Overall, this dissertation focuses on answering the hypotheses presented related to the identification of physical mechanism of fluid injection induced seismicity and the possible application of modern network design and event detection techniques to study induced seismicity. In the process of testing these hypotheses I present earthquake and focal mechanism catalogs for both the FWB and northern Oklahoma regions. Additionally, I present interpretations of the seismically active fault structures and their association with the active

regional stress field in each study region. I also present a fully coupled geomechanical model of the northeast portion of the FWB that links the spatiotemporal history of injection associated pore fluid pressure and poroelastic stress changes in the basin to earthquakes. Lastly, I create two earthquake detection lists using different methodologies for the FWB and northern Oklahoma regions using the *EQcorrscan* template matching and local similarity detection methods, respectively. By combining the analysis of the two study regions and the various earthquake catalogs, models, and event detection lists I create a detailed view of fluid injection induced seismicity occurring within the central United States over varying temporal periods and utilizing differing monitoring network philosophies.

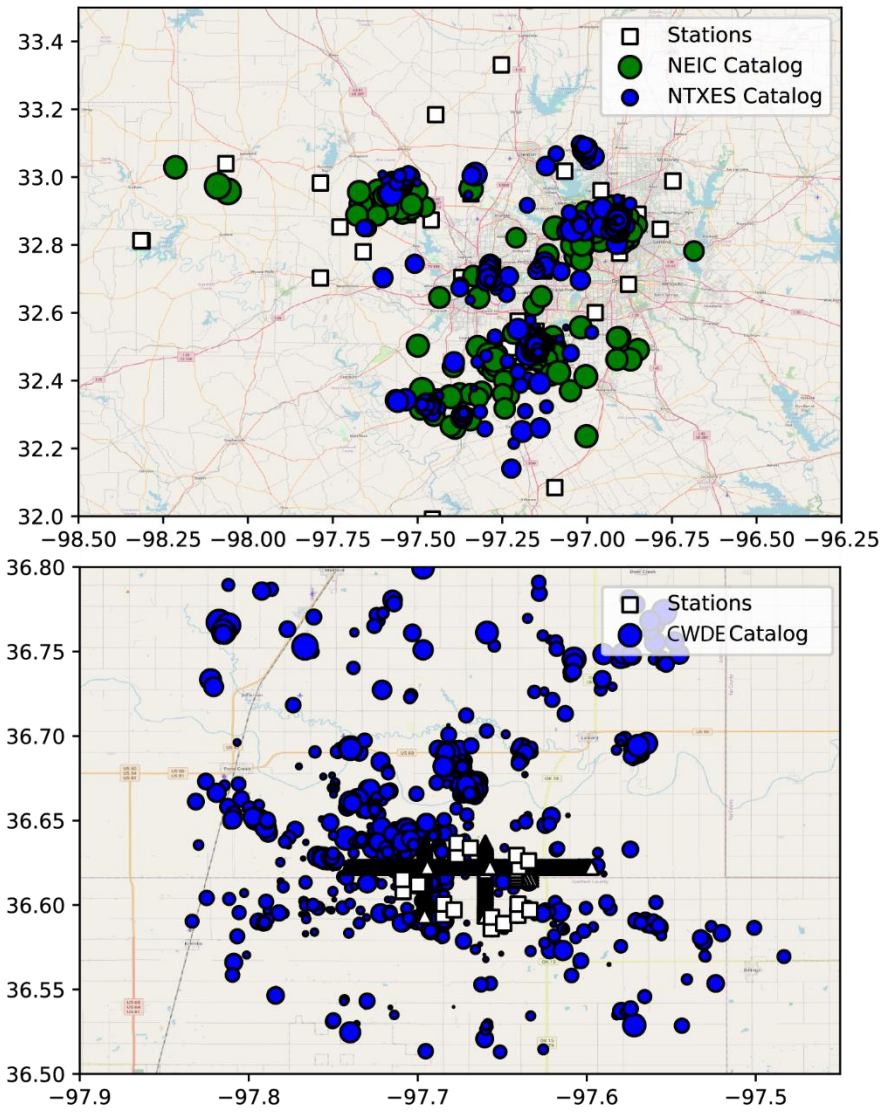

Figure 1.1. (top) Map view of the Fort Worth Basin study area with the recording stations, USGS ComCat catalog, and the NTXES earthquake catalogs shown. (bottom) Map view of the Community Wavefields Demonstration Experiment (CWDE) study area in northern Oklahoma with the recording stations and CWDE earthquake catalog shown.

2. TRACKING INDUCED SEISMICITY IN THE FORT WORTH BASIN: A SUMMARY OF THE 2008-2018 NORTH TEXAS EARTHQUAKE STUDY CATALOG

2.1. Introduction

Starting in late 2008, earthquakes within the Fort Worth Basin (FWB), Texas, contributed to the central United States increased seismicity rates after late-2000's (Frohlich et al., 2010, 2016; Ellsworth, 2013; Weingarten et al., 2015). Studies of individual earthquake sequences in the basin link activity, with varying degrees of certainty, to wastewater injection activities associated with unconventional shale gas development (Frohlich et al., 2010, 2011; Frohlich, 2012; Reiter et al., 2012; Justinic et al., 2013; Hornbach et al., 2015; Scales et al., 2017; Ogwari et al., 2018). Seismogenic faults in the basin are steeply dipping, basement seeded, NE-SW trending normal faults (Magnani et al., 2017; Quinones et al., 2018) (Figure 2.1) and have deformation limited to >300 Ma resolved using formation offset in seismic reflection data (Magnani et al., 2017). Some, but not all, of the larger magnitude earthquakes occur near wastewater disposal wells. Compilations of injection data and estimates of regional pore pressure changes in the FWB (i.e., Gono et al., 2015; Hornbach et al, 2016), however, need to be linked to a more complete documentation in time and space of earthquakes to holistically understand the evolution of the subsurface system. Additionally, the Dallas-Fort Worth metropolitan area (population >6 million) overlies the eastern seismogenic FWB, and a comprehensive catalog of FWB earthquakes provides better data for hazard and risk assessment and regulatory decisions.

The FWB is a foreland basin with a history of oil and gas production activity dating back to the early 20th century (Pollastro et al., 2007) (Figure 2.1). The majority of faults within the basin that have been interpreted from drilling and seismic reflection data have strikes that align well with the strikes of the major basin boundaries (e.g., Ewing, 1990; Pollastro et al., 2007; Magnani et al., 2017; Hennings et al., 2019; Horne et al., 2020). Earthquakes are limited to the northeast portion of the FWB (Figure 2.1). Here, the Barnett Shale formation has served as the primary shale-gas producing unit since 2004 (Pollastro et al., 2007), and wastewater associated with this production is primarily injected into the underlying Ellenburger dolomitic limestone formation (Hornbach et al., 2016). The Ellenburger lies in unconformity atop the crystalline Precambrian basement (Figure 2.2). A complete mapping of basement-seeded faults remains data limited; faults shown in this paper come from recent updated compilation by Hennings et al. (2019).

There are five hypocenter catalogs providing information on earthquakes in the FWB. The catalog of record, the U.S. Advanced National Seismic System Comprehensive Catalog (ComCat), reports mid-magnitude (M_L3) earthquakes consistently through time after 1973, but uncertainty in space can be on the order of 5-15 km. The Frohlich et al. (2016) historic Texas earthquake catalog provides information prior to 1973. Neither of these catalogs contain reliable reported earthquakes in the FWB east of the Bend Arch prior to October 2008. Frohlich (2012) reported small magnitude earthquakes (M<3) in the basin using the Earthscope Transportable Array (TA) from 2009-2011. Between 2008 and 2019, Southern Methodist University (SMU) operated 3 temporary seismic networks deployed over 5 named seismic sequences in the basin (Frohlich et al., 2011; Justinic et al., 2013; DeShon et al., 2018) but focused publication of individual earthquake sequence catalogs over discrete time periods. The North Texas Earthquake Study (NTXES) catalog presented herein, and included within Appendix A.1, reports all

seismicity recorded by the temporary networks operated by SMU during the 2008-2018 period. Finally, beginning in 2017, SMU operations were combined with the Texas Seismic Network (TexNet) such that the NTXES catalog overlaps in time and space with the state-wide, publicly available catalog (Savvaidis et al., 2019).

The NTXES catalog uses a combination of local and regional stations within the basin and a standardized approach to earthquake location and magnitude calculations. The NTXES catalog is comprised of auto-detected and manually reviewed earthquakes located using the GenLoc location algorithm (Pavlis et al., 2004), in conjunction with local and regional 1D velocity models generated using data from well logs collected from within the FWB. I report formal uncertainties for all earthquakes in the catalog. A new regional attenuation curve constrains the local magnitudes reported in the NTXES catalog and is an updated form of the curve provided in Scales et al. (2017) for north Texas. The NTXES catalog is combined with the more temporally complete ComCat to investigate the relationship between earthquakes, faults, and wastewater injection in the FWB and explore magnitude-time relationships along individual faults and within the basin. Finally, I examine the relationship between injected wastewater rates and seismicity and discuss far-field versus near-source triggering effects of fluid injection in the basin, and the possible role fluid injection activities had on the Dallas-Irving sequence, whose primary cause is still under investigation.

2.2. Methodology for the NTXES Catalog

SMU has operated temporary seismic stations in the FWB since 2008 (Frohlich et al., 2011; Justinic et al., 2013), and since 2013 the local networks appear under the auspice of the North Texas Earthquake Study, as summarized in DeShon et al. (2018). Continuous waveform data from all networks are archived without embargo or restriction, including currently operating

stations in near-real-time (see Data and Resources). The networks consist of a mix of short-period, broadband, and strong-motion stations and station locations reflect the complex history of deployment in rapid response mode (DeShon et al., 2018) (Figure 2.1). The resolution in time and space of the resulting NTXES hypocenter catalog reflects this complexity. Early studies using the SMU temporary networks in 2008-2010 used different location methodologies and velocity models (Frohlich et al., 2010,2011; Janska and Eisner, 2012; Justinic et al., 2013; Reiter et al., 2012) than later studies which focused on stations deployed in and after 2013 (Hornbach et al., 2015; Scales et al., 2017; Ogwari et al., 2018; Quinones et al., 2018). In total, there are five well-studied earthquake sequences, here referred to by year and placename of significant first or largest event: 2008 DFW Airport (Frohlich et al., 2010,2011; Janska and Eisner 2012; Rieter et al. 2012; Ogwari et al. 2018), 2009 Cleburne (Justinic et al., 2013), 2013 Azle-Reno (Hornbach et al., 2015; Quinones et al., 2018), 2015 Dallas-Irving (Magnani et al., 2017; Quinones et al., 2018), and 2015 Venus (Magnani et al., 2017; Scales et al., 2017; Quinones et al., 2018). Here, I have joined all data into a single data processing stream to ensure methodological consistency and additionally report all earthquakes, rather than only low uncertainty events associated with specific earthquake sequences.

2.3. Hypocenter Determination

I used the Antelope Environmental Monitoring software and underlying relational database for archiving and analysis of the temporary seismic network data (Boulder Real Time Technologies, brtt.com). Analysis utilized the offline, batch-processing mode and no real-time analysis operations were implemented. The 2008-2011 networks were not telemetered, and while stations post-2013 were, SMU did not have the staff capabilities nor reporting authority to provide real-time earthquake catalogs.

From 2013-present, batch processing 24-hours in arrears includes auto-detection and association of P- and S-wave first arrivals, followed by manual review of associations and raw waveforms to identify small earthquakes. A multi-frequency short-term average over long-term average (STA/LTA) auto-detector (dbdetect) tuned to find impulsive local distance earthquakes feeds into an event associator set to use a spatial grid search method with the iasp91 global velocity model (dbgrassoc). In practice, autodetection and association set to optimize identification across the network can miss emergent or nodal arrivals, trigger incorrectly on a prominent P-to-S converted phase that mixes with first-arriving S on some stations, and do not capture all micro-seismicity ($<M_L 1.0$) associated with swarm activity in some sequences. The network itself exhibits high noise levels inherent to rapid installation within a sedimentary basin and major metropolitan area (discussed in DeShon et al., 2018). Thus, all continuous data is subsequently manually reviewed by a trained analyst to correct autodetections and add additional phase onsets. At this stage, all P-wave first motion data is entered into the database. The analyst assigned phase pick uncertainties associated with these manually reviewed phases are conservatively estimated to be within 0.01-0.04 s for P phase picks and 0.02-0.08 s for S phase picks depending on factors such as the impulsiveness of the phase arrivals and the sampling rates of the observing stations (100 or 200 samples per second).

Event review takes places within the analyst location software (dbloc2), and I use GENLOC location algorithms, which is a modified version of the Gauss-Newton inversion method meant for single event location applications (Pavlis et al., 2004). The GENLOC programs allow for multiple 1D velocity models to be interactively tested resulting in multiple origin locations and times stored for a given event. Reported formal uncertainties include origin time and a 68% confidence error ellipsoid in space and are derive from the covariance matrix in

the inverse solution (Pavlis et al., 2004). The median standard error of observation (sdobs) value, which is defined as the sum of the square of the phase arrival time residuals divided by the number of degrees of freedom, is also stored by origin. For the NTXES catalog, I provided the preferred solution for each event, discussed in the next subsection, and the 68% confidence error ellipsoids are provided as the ellipsoid major axis length and strike, minor axis length, depth axis length, and origin time error (Appendix A.1).

2.4. Velocity Models

The 1D velocity structure of the basin is derived from a combination of available geologic, well log, and reflection data. The FWB stratigraphy summarized in Pollastro et al. (2007) provides the basic geology to inform 1D velocity model design (Figure 2.2). Figure 2.2 is plotted relative to surface, with mean elevation of ~235 meters above sea level. Most significantly, the basin deepens from southwest to the northeast, as reflected in the top of the Ellenburger occurring ~1.3 km below sea level (bsl) in Parker County to over 2.7 km under Dallas County (e.g., Pollastro et al., 2007; Hornbach et al., 2016; Smye et al., 2019) (see Figure 2.1 for placenames). A recent compilation of interpreted well log data across the FWB provides thickness estimates of the Barnett and Ellenburger formations and estimates for the top of the crystalline basement near each earthquake sequence (Smye et al., 2019). I used sonic logs (Figure 2.2b) to constrain P- and S-wave velocity. The Trigg Well (Geotechnical Corporation), located in Tarrant County near the DFW Airport and Dallas-Irving earthquake sequence, and the Briar saltwater disposal (SWD) well, located in Wise near the Azle-Reno sequence, provided sonic logs constraining compressional wave interval velocity through the basin sedimentary units and are in general agreement with one another (Figure 2.2b). The wells also reflect the basin dip; the western Briar Well has a significant velocity jump at 2.2 km and the Trigg Well at ~3 km

below surface reflecting the top of the Ellenburger formation. Dipole sonic logs available at the Bond Ranch SWD well, in western Tarrant County near Azle-Reno, and the A1MD SWD well, near the DFW Airport, suggest Vp/Vs of 1.72 for the Ellenburger and crystalline basement, ranges of 1.82-1.89 through the sedimentary package, and a return to 1.73 in the upper 500 meters. Not many wells drill to top of basement and sonic log data do not indicate a significant velocity contrast between the Ellenburger and crystalline basement. Seismic reflection data in the basin (e.g., Magnani et al., 2017) and the updated FWB stratigraphic model (Smye et al., 2019) confirm an Ellenburger thickness of ~1 km. I used the Briar and Bond Ranch well data to set a 1D model for the Azle region and used the Trigg and A1MD data for DFW Airport, Dallas-Irving, Venus, and Cleburne sequences (Figure 2.2c). Note that previous studies of the Cleburne and DFW Airport relied upon only Trigg well data (Frohlich et al., 2011; Justinic et al. 2013).

Well log data do not constrain the very shallow (<0.5 km) or deep (>5 km) velocity structure required for accurate hypocenter location. Ambient noise analysis of a 10-day deployment of 130 10-Hz vertical component nodes, deployed near Azle (DeShon et al., 2018), yields Rayleigh phase velocities between 0.3 to 0.9 seconds, which are then inverted for 1D Vp and Vs (Sufri et al., 2018). These data constrain the upper 100 m of the Azle 1D velocity model (Figure 2.2b) but were not extrapolated to the other 1D models. TA automated receiver functions place Moho depth between 37-42 km in and near the FWB with a Vp/Vs range of 1.65-1.81 (Data and Resources); I set the Moho depth to 40 km. Frohlich et al. (2011) incorporated a mid-crustal boundary at 18 km to best model arrivals from DFW Airport earthquakes, and regional refraction studies across the Ouachita Thrust Front show a mid-crustal boundary in Laurentia Craton between 20-22 km (Keller and Hatcher, 1999). Velocities provided by Keller and Hatcher (1999) indicate mid-crustal boundaries between 15-25 km, and on FWB stations 18 km best fits

first arrival times. The values are adopted for the mid-crust and lower crust velocities for all 1D models (Figure 2.2d). When an earthquake occurs away from a known, monitored sequence, I adopted the FWB regional velocity model (Figure 2.2d). Models are provided in Appendix A.2 and every earthquake is reported with the associated velocity model in Appendix A.1, available in the electronic supplement.

2.5. Magnitude Determination

Magnitude scaling functions for the FWB and surrounding region are determined using local and regional recordings of earthquakes in the basin between 2013-2018. At close epicentral distances (<100 km) earthquakes are recorded by broadband, short-period and strong motion sensors, while at regional distances (>100 km) the earthquake signals are best recorded by the broadband stations. At very close epicentral distances (<50 km) the dominant recorded phase is the first arriving S-wave; however, at epicentral distances beyond 50 km the Lg-wave begins to dominate the signal (Atkinson and Boore, 2013; Nuttli, 1973). Local magnitude is expressed as:

$$M_L = \log_{10} A(\Delta) - \log_{10} A_0(\Delta) + c \quad (2.1)$$

where $\log_{10} A(\Delta)$ is the base-10 logarithm of the peak amplitude (in millimeters) on a Wood–Anderson seismometer measured at some epicentral distance Δ (in kilometers), and c is a station correction term that is not applied in this study (Richter, 1935). The $\log_{10} A_0(\Delta)$ term is a distance scaling factor that is determined by constraining the zero-point of the magnitude scale to a hypothetical Wood-Anderson instrument. For instance, at 100 km from the epicenter the peak amplitude of a ML 3.0 earthquake is equal to 1 mm as defined by Richter (1935). With a c value of 0, the distance scaling factor can thus be expressed as:

$$\log_{10} A_0(100) = \log_{10} A(100) - 3 = \log_{10} 1 - 3 \quad (2.2)$$

The $\log_{10} A_0(\Delta)$ term is empirically derived by first convolving instrument corrected waveforms with a Wood-Anderson instrument response. I then sampled events with at least one recording station at an epicentral distance of approximately 100 km, which was then used as a normalization station for that earthquake. Peak amplitudes were derived from the greater of the two horizontal component waveforms bandpass filtered between 0.1-5.0 Hz (Figure 2.3), following the original practice described in Richter (1935) and adopted by USGS for computation of M_L (Patton et al., 2016). Normalization of all stations for each event to the recording at ~100 km conditions each earthquake to ML 3.0. Note that a previous local magnitude scale derived using FWB data calibrated M_L to the USGS ComCat reported mb_Lg following the method of Walter et al. (2016) (Scales et al., 2017). The Scales et al. (2017) relation was previously adopted as M_L for the Texas Seismic Network but has since been updated as well (Savvaidis et al., 2019).

Here, the derived an attenuation curve uses recordings of earthquakes reported in the ComCat, following Scales et al. (2017), but normalized as described above. Events reported in the ComCat exhibit good signal-to-noise at regional broadband stations (Figure 2.1c) and at broadband and strong motion stations within the basin. The calculation additionally uses recordings from the TA between 2008 and 2011. The initial earthquake set (black circles, Figure 2.3) yields primarily regional distance data out to 400 km. The best fit attenuation curve (gray line, Figure 2.3) models amplitude of first arriving S, transition to Lg. This curve-fit represents the attenuation of the Lg waves and could also be considered an mb_Lg magnitude equivalent.

This station-event dataset is identical to Scales et al. (2017), but the change in normalization significantly reduces scatter in the amplitudes at individual stations (Scales et al., 2017) and matches ComCat mb_Lg without need of additional correction. The resulting ML relationship is:

$$ML_{2019A} = \log_{10} A_0(\Delta) - 1.19 \log_{10}(\Delta) - 0.6 \quad (2.3)$$

The NTXES catalog contains many very small earthquakes that were not recorded at 100 km or on the broadband and strong motion sensors originally analyzed. The ML_Scales and ML_{2019a} relationship significantly overestimated peak amplitudes for local stations (<50 km). Therefore, I normalized the short epicentral distance peak amplitudes (light gray symbols, Figure 2.3) using stations at 50 km distance and then adjusted the amplitude values to the zero-point based on the 100 km normalization distance data. The resulting ML is:

$$ML_{2019b} = \log_{10} A_0(\Delta) - 1.9 \log_{10}(\Delta) + 0.6 \quad (2.4)$$

This attenuation curve calculates ML that match the mb_Lg well for earthquakes at local distances (<50 km) but overestimates the magnitudes of earthquakes at regional distances (Figure 2.3). The scatter in the plot is attributed partly to stations' site effect and radiation pattern.

The NTXES catalog reports a single magnitude per earthquake calculated using the Antelope software magnitude calculator dbevproc. Any event reported within the NTXES catalog recorded using only stations within the 50 km epicentral distance limit uses the ML_{2019b} attenuation curve function, which is included as a modification to the dbevproc parameter file. Meanwhile, if an event recorded within the FWB uses many regional stations at distances

exceeding 50 km, then the ML_{2019a} attenuation curve function is applied. However, no events reported in the 2008-2018 NTXES catalog use regional phases because the 1D velocity models are designed for local network data, and hence regional phases are not integrated into the Antelope database, even for larger earthquakes. In practice all M_L reported in the NTXES catalog through 2019 reflect ML_{2019b}. Uncertainty is estimated to be on the order of 0.1 – 0.3 units. Figure 2.4 shows the cross-plot between ML_{2019b} and mb_lg for earthquakes reported in the NTXES and ComCat catalogs, respectively.

2.6. Earthquake Catalog

The seismicity reported in the NTXES catalog describes individual earthquake sequences along linear features identified as faults and contains individual earthquakes scattered in time that are not easily ascribed to known faults (Figure 2.5). The catalog describes two separate time periods of seismic monitoring activity: 2008-2010 and post-2013 (Figure 2.5). In the NTXES catalog, I identified 9 active earthquake sequences on discrete faults described by their location and year of initial activity here: DFW Airport (2008), Cleburne (2009), Azle-Reno (2013), Dallas-Irving (2015), Venus (2015), Haslet (2015), Lake Lewisville (2017), Fort Worth (2017), and west Cleburne (2018). Of these, Lake Lewisville, Fort Worth and west Cleburne had not been previously reported and had only 1-2 monitoring stations within a 10 km hypocentral distance. These three sequences are shown in cross-section in Figure 2.6 but have significant depth uncertainties compared to the well-recorded Azle-Reno, Dallas-Irving, and Venus sequences. Figure 2.7 shows the formal uncertainties for the NTXES catalog, subdivided by the three significant post-2013 event sequences and all earthquakes located using the regional velocity model. Taking the entire dataset, median values for major, minor and depth axes are <0.4 km and median origin time error is 0.04 s. The residual measure, sdobs, also has a median

of 0.04 seconds. Individual event uncertainties can range higher, however, and I provide formal error estimates for each event in Appendix A.1.

The majority of earthquakes in the FWB are occurring within the Precambrian crystalline basement (Figure 2.6b-g). However, a portion of earthquakes within the Azle-Reno and regionally located sequences locate within the Ellenburger formation (Figure 2.6b,e,f,g). The shallower Azle-Reno events are associated with an antithetic feature near the main fault (Hornbach et al., 2015), while the shallower events in the regionally located sequences are likely an effect of larger depth uncertainties due to a lack of close hypocentral distance stations.

Each of the earthquake sequences within the FWB exhibit swarm-like behavior rather than resembling main shock-aftershock sequences. Figure 2.5 shows the magnitude versus time distribution of all earthquakes recorded within the FWB separated by reporting catalog. The characteristic distribution of seismicity over time associated with each individual sequence in the FWB is a relatively short period (6-12 months) of peak seismicity followed by a steep decline in subsequent seismicity. However, overall basin-wide seismicity rates have remained steady since the onset of recorded seismicity in the FWB in 2008. Thus, the gap in recorded seismicity within the NTXES catalog from 2010-2013 is a gap in local seismic monitoring capabilities rather than seismic activity.

2.7. Magnitude Distribution

All NTXES catalog magnitudes are calculated using the ML_{2019b} attenuation curve function (Figure 2.3). The overall magnitude range of earthquakes within the NTXES catalog is M_L-1.0 to M_L4.0, although the magnitude range of each individual sequence varies. The differences in the number of stations, station geometry, and overall noise levels across the

sequences described in the NTXES catalog have led to large variations in the degree of catalog completeness across the nine sequences in the FWB.

There is spatial variation in the magnitude of completeness (Mc) and b-values of the sequences described within the NTXES catalog. The NTXES catalog is divided into four subgroups; the three significant post-2013 sequences being the Azle-Reno, Dallas-Irving, and Venus sequences, and the sequences located using a regional velocity model hereafter referred to as the regional sequences. Mc and b-values are calculated for each subgroup using the 90% goodness-of-fit method (Wiemer and Wyss, 2000), and the maximum likelihood estimation method, respectively (Bender, 1983). There are a wide range in Mc values across the subgroups from Mc=0.0 for the Venus subgroup to Mc=2.1 for the regional events subgroup (Figure 2.8). The higher Mc value for the regional events can be attributed to a lack of local stations to monitor the small magnitude events. Meanwhile, the variation in Mc values across the three locally monitored sequences in the NTXES catalog can be attributed to a non-optimal initial network geometry in the case of the Azle-Reno sequence, and to elevated noise levels in the Dallas-Irving sequence, which is embedded in the metroplex (DeShon et al., 2018). The b-values also vary across the subgroups from 0.67 for the Dallas-Irving subgroup to 1.01 for the Azle subgroup (Figure 2.8). The lower b-values calculated for the Dallas-Irving and Venus subgroups are likely due to a variety of factors such as: Mc uncertainties (Woessner and Wiemer, 2005), missing early lower magnitude events prior to instrument deployment (Scales et al., 2017), or noise level issues (DeShon et al., 2018). Additionally, using the same methodology, we calculated that the 2008-2018 ComCat catalog of FWB earthquakes has a Mc value of 2.6 and b-value of 1.25 (Figure 2.8). Lastly, the b-value for the NTXES catalog use the Mc value from the regional events subgroup (2.1) and have a value of 0.74. To assess whether there were temporal

variations in the Mc and b-values associated with each sub-group relating to either the changing instrumentation or changing stress state, I ran the Mc calculations using a variety of temporal sub-groups for each area (Figure 2.9). Additionally, to provide a secondary check on the associated Mc and b-values of each sub-group I calculated the b-value associated with each sub-group using a range of set Mc values (Figure 2.10). The first test showed that while there is some temporal variability in the Mc and b-values of each sub-group, these variations do not appear to be related to the instrumentation or regional stress state of each sequence but rather are more related to the increasing completeness of the catalog. The second test showed that the estimates of the Mc value for each sub-group did tend to align with a range of Mc values for which the associated b-value was stable following the criteria for estimating Mc published in Cao and Gao (2002).

2.8. Location Changes from Prior Catalogs

The seismicity reported within the NTXES catalog for the DFW airport, Cleburne, Azle-Reno, Dallas-Irving, and Venus areas has been presented and discussed in prior publications but now have some differences in earthquake locations to the NTXES catalog presented here. The NTXES catalog locations are different than in prior publications due to the updating of velocity models used for earthquake location (Figure 2.2). Previous publications presenting the original velocity models used to calculate earthquake locations were based on then available well log and geologic data. The updated velocity models have been supplemented with newly available sonic log, seismic reflection, ambient noise tomography, and geologic data as described in the Velocity Models section. Overall, there is little change in the earthquake hypocenter locations with median changes (with 50% confidence error values) of 0.00±0.07, -0.01±0.07, and 0.00±0.19 km for the latitude, longitude, and depth respectively from the prior published catalogs to those

presented here. Additionally, origin time differ only slightly with a median change of -0.01±0.03 s (Figure 2.11). Therefore, while the earthquake locations have slightly changed with the updating of the velocity models, the fault structures described by the distributions of the earthquake locations at each sequence site are similar. Thus, previous interpretations of fault geometries and earthquake location distributions remain valid.

2.9. Earthquakes and Faults

Most active faults in the FWB are NE-SW trending normal faults which are concentrated in the northeast portion of the basin (Figure 2.1b). Prior studies focusing on the three most significant post-2013 FWB sequences have utilized the NTXES catalog to interpret fault geometries and deformation histories of the active faults in the FWB using focal mechanism (Quinones et al., 2018) and seismic reflection data (Magnani et al., 2017). The focal mechanisms generated from the NTXES catalog described the source faults of each of these sequences as steeply dipping normal faults with strikes of ~40° and dips of between 56°-70° (Quinones et al., 2018). Meanwhile, seismic reflection data collected across the Venus and Dallas-Irving sequence sites revealed a lack of vertical displacement on the faults at both sites in rocks younger then ~310 Ma (Magnani et al., 2017). This implied that the faults at both sites remained inactive since the Pennsylvanian. Subsequently, additional proprietary seismic reflection data collected across the basin revealed more wide-spread NE-SW trending faulting in the northeast portion of the FWB. Also, the majority of NE-SW trending faults in the FWB are considered optimally oriented for failure with high slip potentials within the local and regional stress fields described by the focal mechanisms and borehole breakout data collected from the basin (Quinones et al., 2018; Hennings et al., 2019). In each study the NTXES catalog was essential in providing proper constraints and interpretations of the resulting imaged fault structures in the FWB.

The NTXES catalog contains a record of seismicity occurring within the previously undocumented Lake Lewisville (2017), Fort Worth (2017), and west Cleburne (2018) sequences. The earthquakes within these sequences are located using the regional velocity model with data collected by a combination of TexNet and SMU operated stations, although no dedicated local networks have been installed at any of these sequence sites. Thus, fewer earthquakes have been detected within these sequences and those that have been located have higher associated depth uncertainties. Because of these issues, I could not provide the same degree of fault interpretation for these three sequences in comparison to the other post-2013 sequences. The Lake Lewisville sequence consists of 17 earthquakes with depths ranging from 2-9.5 km which appear to occur along a steeply dipping NE-SW trending fault plane (Figure 2.6e). The Fort Worth sequence consists of only 9 detected earthquakes ranging in depth from 2-7 km (Figure 2.6f). We have not provided a fault interpretation for the Fort Worth sequence due to the lack of associated hypocenter locations. The west Cleburne sequence is the most recent to become active, however, its associated earthquake count has already surpassed those of the Lake Lewisville and Fort Worth sequences. The west Cleburne sequence earthquakes have the highest location uncertainty values in the NTXES catalog due a sizable network azimuthal gap and the lack of local stations for depth control. The earthquakes in west Cleburne range in depths from 1-5 km and appear to describe a steeply dipping N-S trending fault similar in orientation to the fault described by the original Cleburne sequence (Figure 2.6g, Justinic et al., 2013). This N-S trending fault interpretation means this fault would not be optimally oriented for failure within the previously reported FWB stress regimes (Lund Snee and Zoback, 2016; Quinones et al., 2018); however, seismic reflection and well head data interpretation also point to a N-S trending fault at this location (Hennings et al., 2019).

2.10. Earthquakes and Injection Data

The seismicity occurring within the FWB is part of the larger trend of increasing amounts of induced seismicity within the central United States (US) which has been associated with fluid injection activities. Pore pressure diffusion associated with fluid injection activities is hypothesized to be the primary mechanism driving induced seismicity within the FWB (Frohlich et al., 2016; Hornbach et al., 2016) and throughout the central US (e.g., Keranan and Weingarten, 2018). Monthly volumes of fluids injected into the Ellenburger formation, the main disposal unit in the basin, by SWD wells are reported by the Texas Railroad Commission (TRC) and can be accessed electronically using their public database. Over the time period of Oct. 2005-Oct. 2017 over 2 billion US barrels of fluids from 179 SWD wells were injected into the Ellenburger formation (Figure 2.12). When I examined an interpolated surface describing the cumulative volumes of injected fluids from 2005-2017, I observed that the northeast portion of the FWB is where both the majority of injection activities and seismicity is occurring within the basin (Figure 2.13). In fact, with the exceptions of the Dallas-Irving and Lake Lewisville sequences, the majority of seismicity within the FWB is occurring within 15 km of at least one injection well. The spatial proximity of these near well sequences, along with the strong temporal correlation between the onset of seismicity and increasing injection rates within the FWB (Figure 2.13, inset), suggests that pore pressure diffusion is the main driving force for induced seismicity at these sequence sites. However, injection rates have decreased in recent years from their peak levels in 2014, mainly due to economic reasons, which does appear to coincide with lowering rates of seismicity across the FWB. Previous studies using the NTXES catalog data focusing on these near well sequences have found that pore pressure changes associated with injection activities are significant and are the primary mechanism driving seismicity at these sites

(Frohlich et al., 2011; Hornbach et al., 2015, 2016; Ogwari et al., 2018; Quinones et al., 2018; Scales et al., 2017). However, stress changes associated with pore pressure diffusion are often limited to distances close to wells (<15 km) (Goebel et al., 2017; Segall and Lu, 2015) leaving the question for what the main mechanisms driving seismicity at sites that are at far distances from injection wells.

2.11. Far-field vs. Near-Source Triggering

While pore pressure changes caused by fluid injection activities are the dominant stress change effect at near well distances, modeling results have shown that at farther distances from injection wells (>15 km) poroelastic stress changes dominate. Recent studies on stress changes associated with injection activities have focused on not only understanding direct pore pressure changes, but also on understanding the far-field effects of poroelastic stress changes (Chang and Segall, 2016; Goebel et al., 2017; Segall and Lu, 2015). In the FWB, two sequences occur away from injection wells: the Dallas-Irving and Lake Lewisville sequences (Figure 2.12).

Results of injection related stress change modeling predict a crossover distance at which poroelastic stress effects become dominant over direct pore pressure stress changes (Goebel et al., 2017; Segall and Lu, 2015). However, this crossover distance is highly variable, relying on factors such as the properties of the injection unit, injection rate, and the duration of injection activities. Prior studies have sought to model pore pressure stress changes within the FWB focusing on the basin-wide effects of injection activities (Gono et al., 2015; Hornbach et al., 2016; Zhai and Shirzaei, 2018; Gao et al., 2021) and the localized stress changes associated with injection activities at the DFW airport (Ogwari et al., 2018). Hornbach et al. (2016) found the Ellenburger to be over pressured by about 1.7-4.5 MPa at injection well sites in northeast Johnson County, and Zhai and Shirzaei (2018) calculated overpressure within the Ellenburger to

be ~2 MPa in that same area. Ogwari et al. (2018) also found that injection activities increased pore fluid pressure within both the Ellenburger and basement formations in the DFW airport area. The Gao et al. (2021) hydrogeologic model is the most recent of these model results and also found peak pore fluid pressure changes within the Ellenburger to exceed 2 MPa. In these studies, stress changes associated with direct pore pressure effects are highly concentrated at close distances to the wells. Thus, it's believed that poroelastic rather than pore pressure stress changes are the primary driving mechanism of seismicity sequences at far distance sites. However, while poroelastic stress changes are dominant over pore pressure stress changes at far distances, the actual magnitude of the poroelastic stress changes is still lower than the near well pore pressure effects (Segall and Lu, 2015). This leads to a larger question, still remaining to be resolved in the FWB, would poroelastic stress changes alone be large enough to have induced slip on the far distance sequences? Previous studies have attempted to calculate the slip probability and stress change necessary to induce slip of the Dallas-Irving sequence fault (Hennings et al., 2021; Quinones et al., 2018). Both studies have determined that the Dallas-Irving fault is optimally oriented for failure within the given stress field (3.48±2.39 MPa), but it's unclear whether poroelastic stress changes alone would be enough to induce slip on the fault. It should be noted that the prior estimate of stress changes needed to have induced slip on the Dallas-Irving came from an estimate utilizing a stress tensor created using focal mechanism stress inversion data and other estimates of needed to stress changes to induce slip are of lower magnitudes in the range of 0.05-0.31 MPa when analyzing modeled pore fluid pressure changes at the fault locations at the onset of seismicity (Quinones et al., 2018; Hennings et al., 2021). The formal focal mechanism stress inversion result was calculated using the full range of fault plane solutions from the focal mechanism catalog of the Dallas-Irving fault. However, when using only

the fault plane orientation value for the Dallas-Irving fault determined using the combined earthquake distribution, focal mechanism and seismic reflection data, I calculated a stress change of 0.26 MPa would be needed to have induced slip on the Dallas-Irving fault (Figure 2.14). This value is of a lower magnitude than the previous focal mechanism derived estimate due to the exclusion of those fault plane orientations that least matched the active local stress field; however, given the incorporation of all separate fault interpretation data sets this value better estimates the needed stress change to induce slip on the fault using the calculated stress tensor.

2.12. North Texas

The FWB seismicity shares many characteristics with other induced seismicity sites occurring throughout the central US such as: primarily being concentrated near injection wells, occurring within the basement formations, and having a strong temporal correlation with increasing injection rates. Numerous catalogs of induced seismicity throughout the central US report that the majority of seismicity is occurring along faults residing within the pre-Cambrian basement which typically underlies the main fluid disposal unit in the region. This pattern in induced earthquake depths has been observed in the FWB, Guy-Greenbriar (Horton, 2012), Raton Basin (Rubinstein et al., 2014), Oklahoma (Keranan et al., 2014), southern Kansas (Rubinstein et al., 2018), and Delaware Basin sequence sites. This is not to say that no seismicity occurs within the units above the basement; seismicity was also recorded within the fluid disposal unit at each of the aforementioned sequence sites. These earthquake depth distributions imply that the active faults within the basement formations are either hydraulically conductive or connected to the fluid disposal formations (Chang and Segall, 2015).

Seismic reflection data collected across the Dallas-Irving and Venus regions shows that the faults at each site stretch into the overlying units above the basement (Magnani et al., 2017).

This means the faults themselves could act as the connection between the fluid disposal and basement units allowing for the transfer of pore pressure between them. Local distance pore pressure modeling work focusing on the Azle-Reno and Venus sites has been completed (Haddad and Eichhubl, 2020; Park et al., 2021); however, such modeling efforts are not currently underway for the Dallas-Irving site due to its far distance from injection wells. Presently, all measurements of pore pressure and poroelastic stress changes affecting the Dallas-Irving site come from basin-wide modeling efforts which calculated very little to no stress changes in the Dallas-Irving area (e.g., Zhai and Shirzaei, 2018; Gao et al., 2021). The Dallas-Irving sequence, still the most enigmatic of the sequences, generated significant felt earthquakes within the NTXES catalog, and determining the main driving mechanism behind its seismicity will require a better understanding of how pore pressure, poroelasticity, and the injected fluids flow and diffuse throughout the FWB.

2.13. Summary

The NTXES catalog represents the most complete record of seismicity occurring within the FWB. All catalog earthquake locations are manually reviewed and calculated using the GENLOC location algorithm within the Antelope database software system. The earthquake hypocenter locations along with their 68% confidence error ellipsoid information are reported within the catalog (Appendix A.1). The 1D velocity models used for locating the FWB earthquakes were generated using a combination of geologic, well log, ambient noise, receiver function, and seismic reflection data collected from across the basin. All magnitudes reported in the NTXES catalog are local magnitudes calculated using new specialized regional attenuation curve functions for earthquakes located using either local or regional distance station data. As a whole, the NTXES catalog earthquakes have low location uncertainties due to the majority of

events being located by dedicated local seismic networks at close epicentral distances with good azimuthal coverage. In the NTXES catalog I identified 9 separate earthquake sequences occurring along discrete steeply dipping NE-SW trending normal faults located primarily within the pre-Cambrian basement formation. The Mc of the NTXES catalog varies across the sequences due to differences in station density and network geometry, however the overall Mc of the catalog is lower than that of other seismicity catalogs in the FWB such as the ComCat catalog. Overall, seismicity in the FWB does have a strong spatial and temporal correlation with fluid injection activities with the majority of seismicity occurring within 15 km of SWD wells. The main exceptions to this are the Dallas-Irving and Lake Lewisville sequences which have no SWD wells within 15 km. This means that far-field rather than near-source stress changes may contribute to driving seismicity at either sequence site. Future work involving the NTXES catalog will focus more on the modeling of geomechanical stress changes associated with fluid injection activities to discern the main mechanisms driving seismicity at both the near and far from well distance sequence sites.

2.14. Data and Resources

All seismic data used in this study were collected as part of the North Texas Earthquake Study projects focusing on the study of seismicity occurring within the northeastern portion of the Fort Worth Basin. These projects were conducted by Southern Methodist University using a combination of SMU, USGS, IRIS PASSCAL, and TexNet instruments. The data used in this study can be obtained from the IRIS Data Management Center at www.iris.edu under the Federated Digital Seismic Network codes NQ, ZW, 4F, and TX (last accessed February 2019). Transportable Array receiver function information can be accessed using the IRIS Earthscope Automated Receiver Survey (EARS) data services product at

https://doi.org/10.17611/DP/EARS.1. Injection volume information for saltwater disposal wells in the Fort Worth Basin can be obtained from the Texas Railroad Commission's online public database at webapps.rrc.texas.gov/H10/h10PublicMain.do (last accessed January 2019). The TexNet earthquake catalog information can be obtained from their public online website at www.beg.utexas.edu/texnet-cisr/texnet (last accessed February 2019). The USGS Advanced National Seismic System Comprehensive Catalog information can be obtained from the public online website at www.earthquake.usgs.gov/earthquakes/search/.

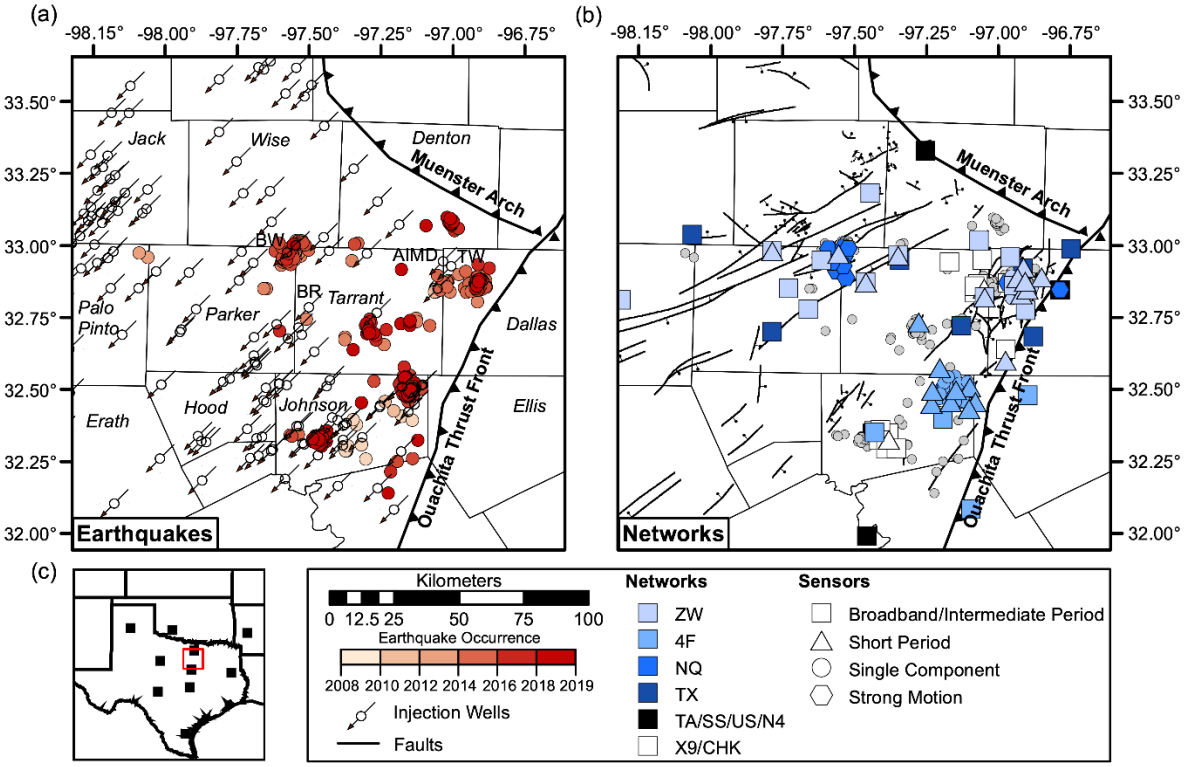

Figure 2.1. (a) Map view showing the locations of the NTXES earthquakes as circles shaded by the time of their occurrence along with the locations of wastewater wells (arrows) in the basin that were active during the period of observation. County names (italics) and important well locations such as the Bond Ranch (BR), Briar Well (BW), Trigg Well (TW), and A1MD well are also labeled. (b) Map view showing the locations of all stations that were used to locate the NTXES earthquakes shaded by their network codes and whose symbols represent the station's sensor type. The locations of the NTXES earthquakes (light gray circles) are also shown. Faults interpreted from proprietary seismic reflection data (Hennings et al., 2019). (c) General map view showing the locations of regional US and TA stations used to locate some NTXES catalog earthquakes along with the highlighted study area (box).

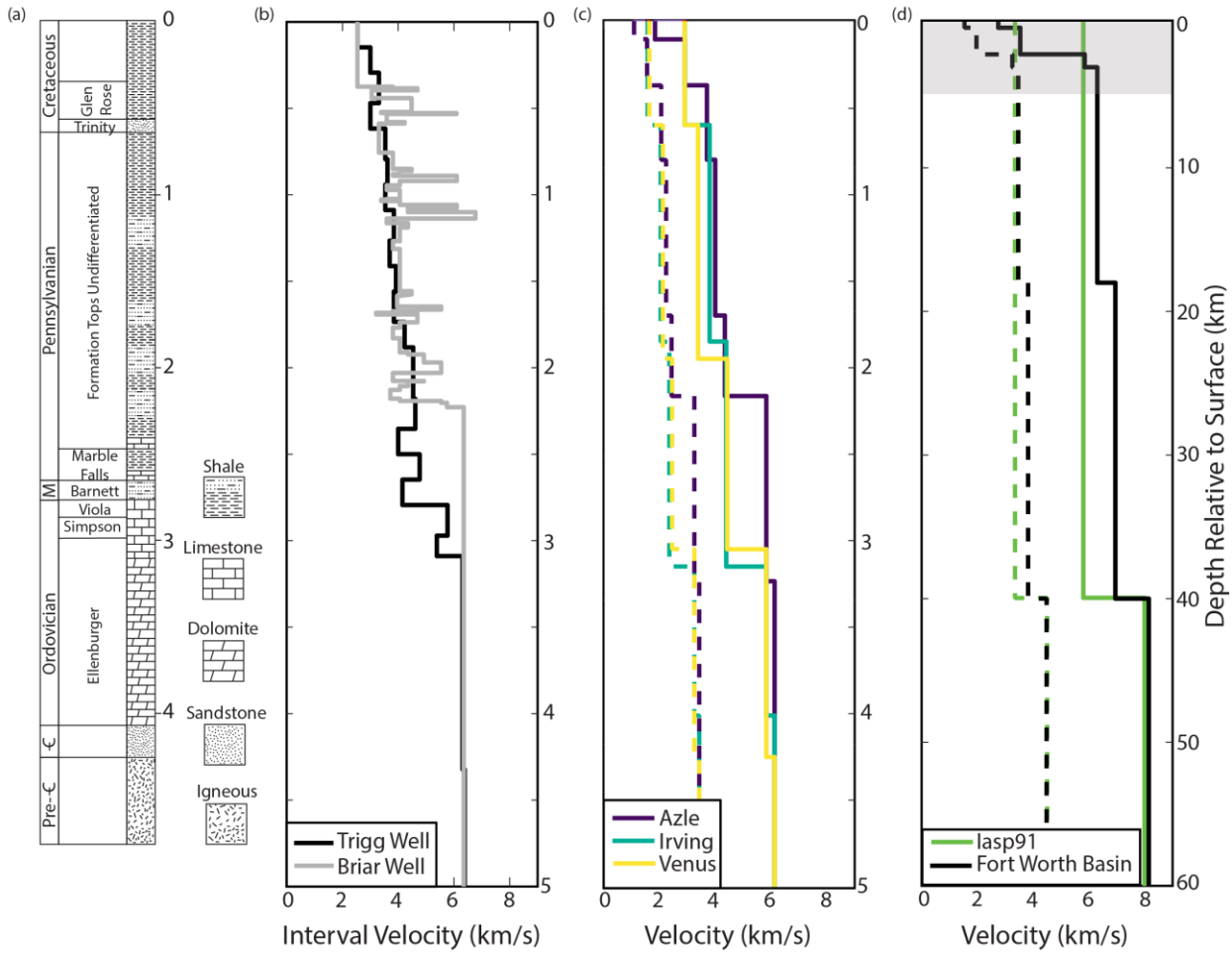

Figure 2.2. (a) Stratigraphic column created using data collected from the Trigg Well site. (b) Interval velocity models created using data collected from the Trigg and Briar Well sites. (c) 1D local P- (solid lines) and S-wave (dashed lines) velocity models used to locate earthquakes within the Azle, Dallas-Irving, and Venus sequences. (d) 1D regional P- (solid lines) and S-wave (dashed lines) velocity models used to locate earthquakes within the Fort Worth Basin which occur outside the three previously mentioned sequences. The upper 5 km of the regional velocity models which is similar to the local 1D velocity models is highlighted (gray area).

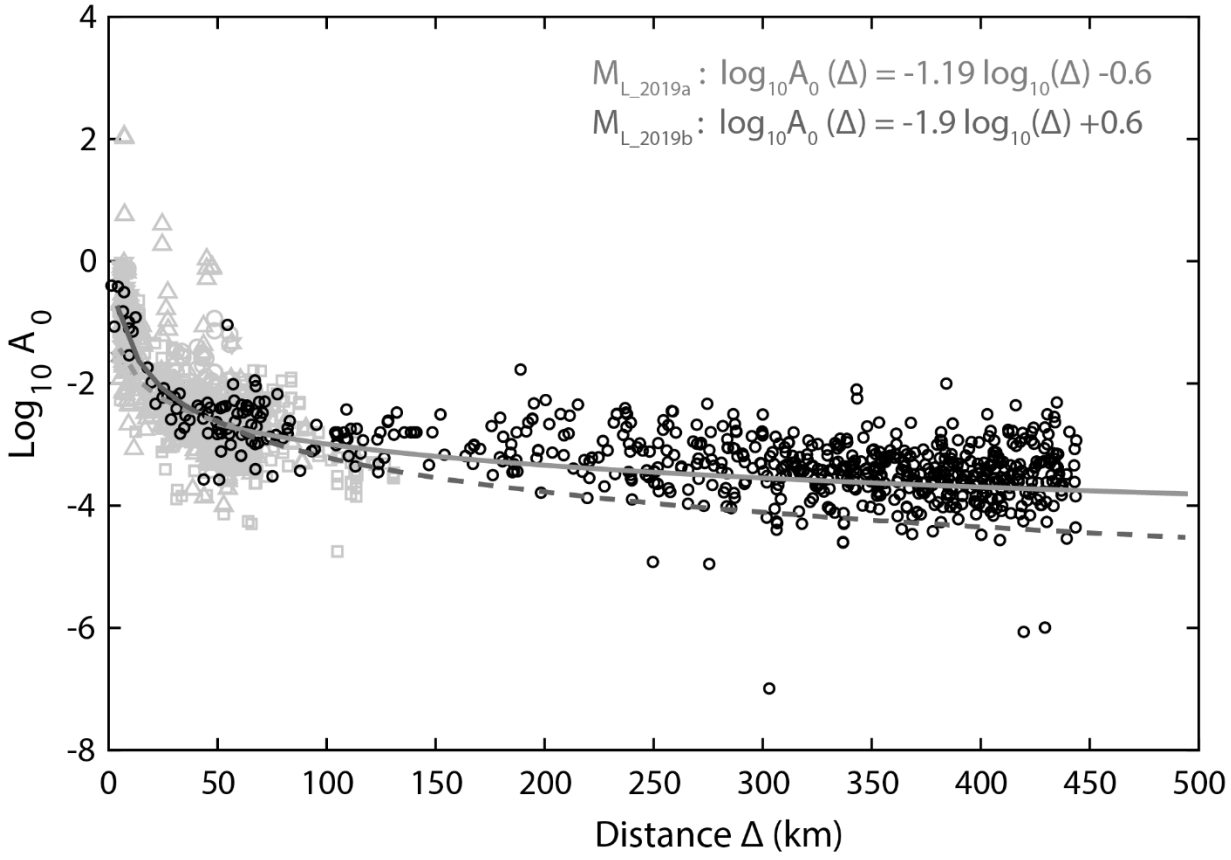

Figure 2.3. Attenuation curves created for the Fort Worth Basin. Light gray symbols represent peak amplitudes normalized to a station located 50 km from the epicenter; symbol shape follows Figure 1. Black circles represent peak amplitude normalized to a station located 100 km for earthquakes reported in ComCat. The dark gray line best fits small magnitude earthquakes recorded by the SMU networks, while the gray line best fits regional broadband data. Hence, we adopt ML_2019b attenuation relation for data recorded at <50 km and the ML_2019a for data at >50 km (solid portions of lines).

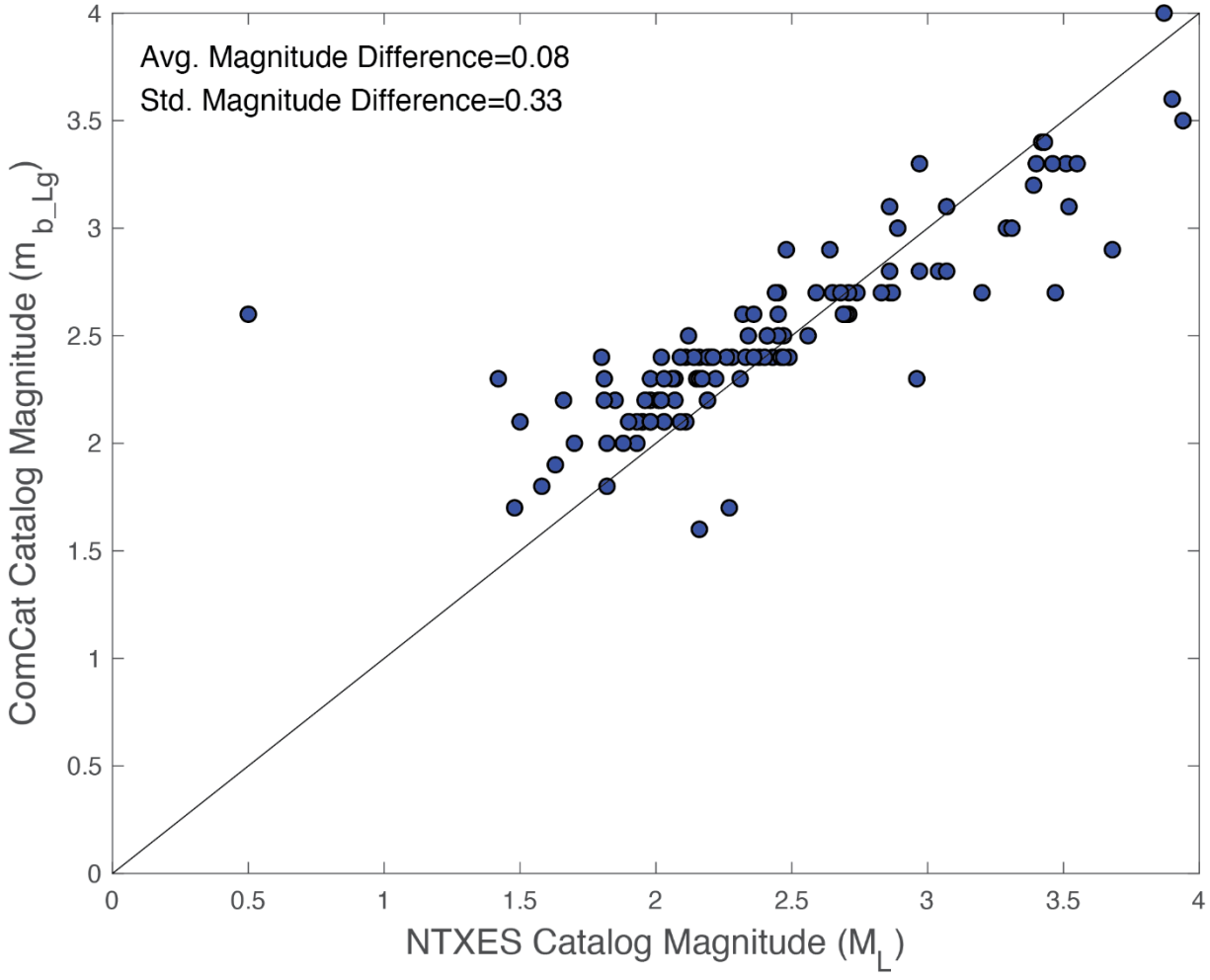

Figure 2.4. Comparison of magnitudes for earthquakes recorded in both the NTXES (ML) and ComCat (Mb_Lg) catalogs along with the average (Avg.) and standard deviation (Std.) magnitude difference values shown.

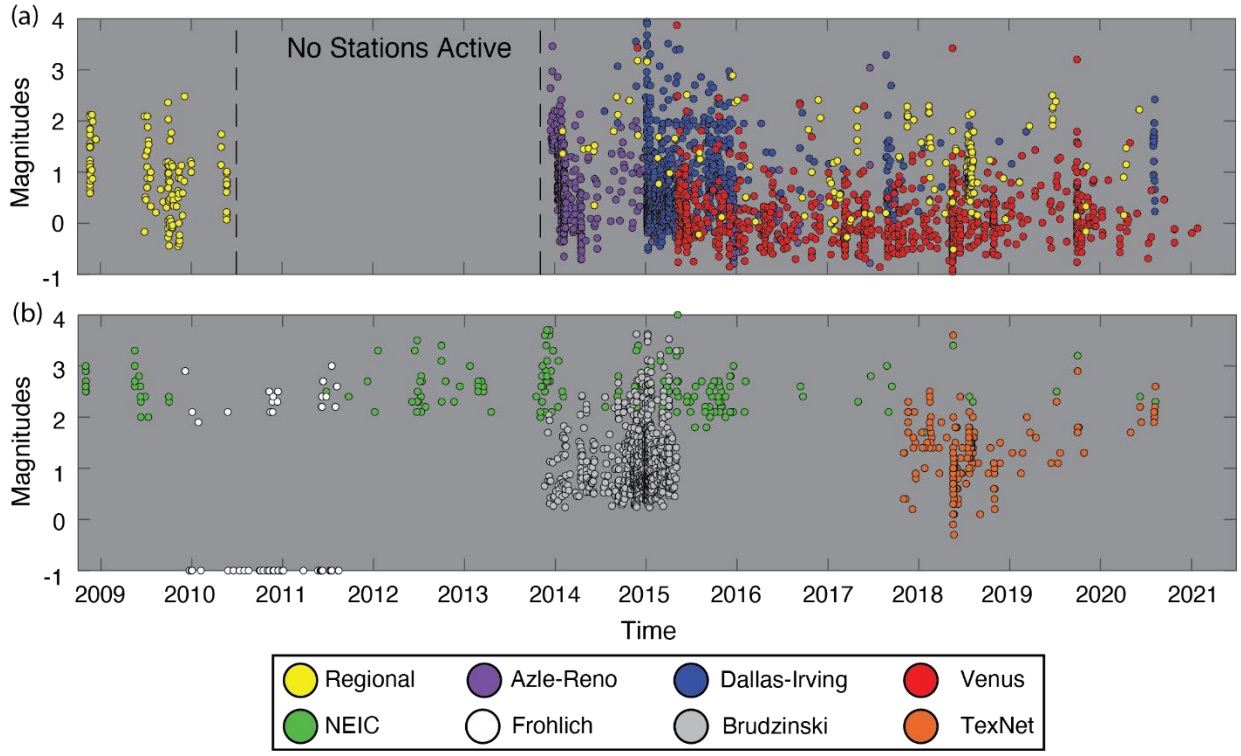

Figure 2.5. (a) Magnitude vs. time plot of the NTXES catalog separated into the Azle, Dallas-Irving, Venus, and Regional sub-groups (dashed lines indicate time period where there were no active stations within the Fort Worth Basin). (b) Magnitude vs. time plot of the earthquakes located within the Fort Worth Basin from the Frohlich 2012, USGS ANSS ComCat (NEIC), Brudzinski template matching and TexNet earthquake catalogs.

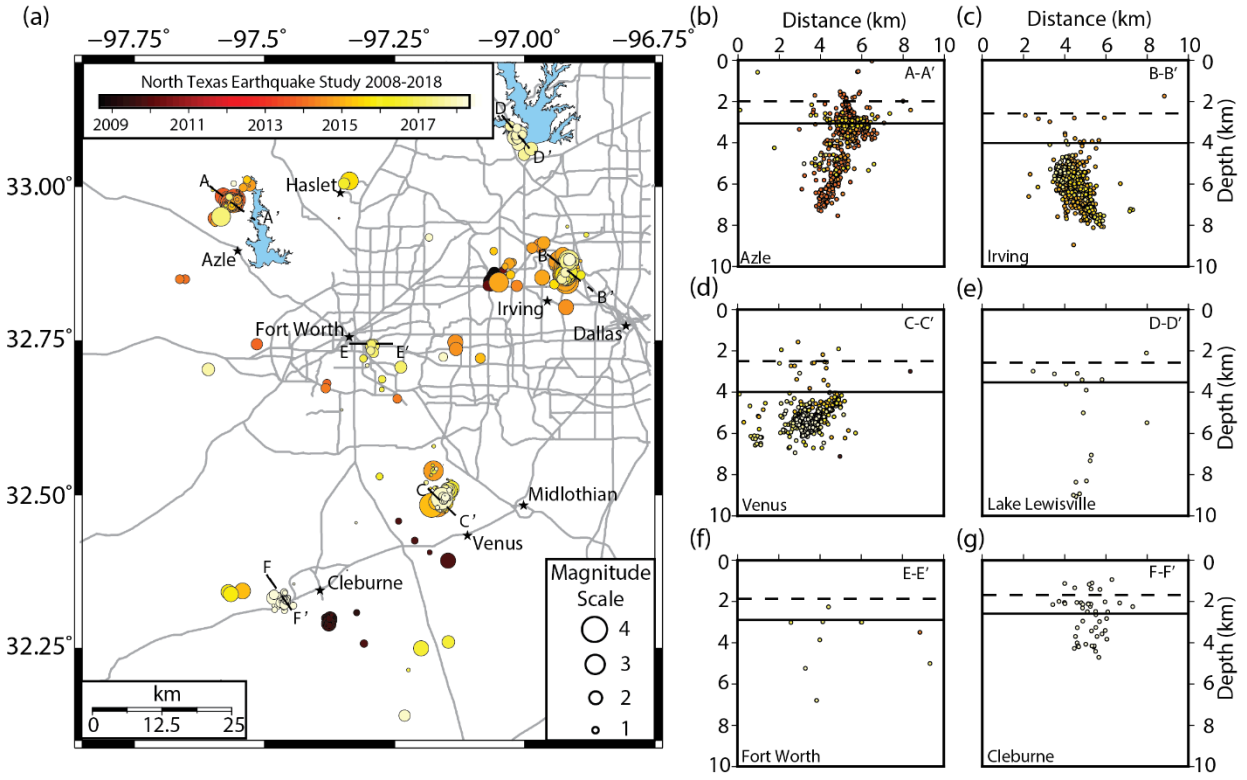

Figure 2.6. (a) Map view of the NTXES study area showing the locations of all earthquakes in the catalog shaded by their time of occurrence and scaled by their magnitude. Also shown are the major roads and highways (gray lines) in the region. The lettered dashed lines represent profile lines used to create the separate cross-sectional views for each named sequence. (b-g) Cross sectional views of each named sequence site using the same occurrence time scale. Also shown are the depths of the top of the Ellenburger (dashed lines) and basement (solid lines) at each sequence site.

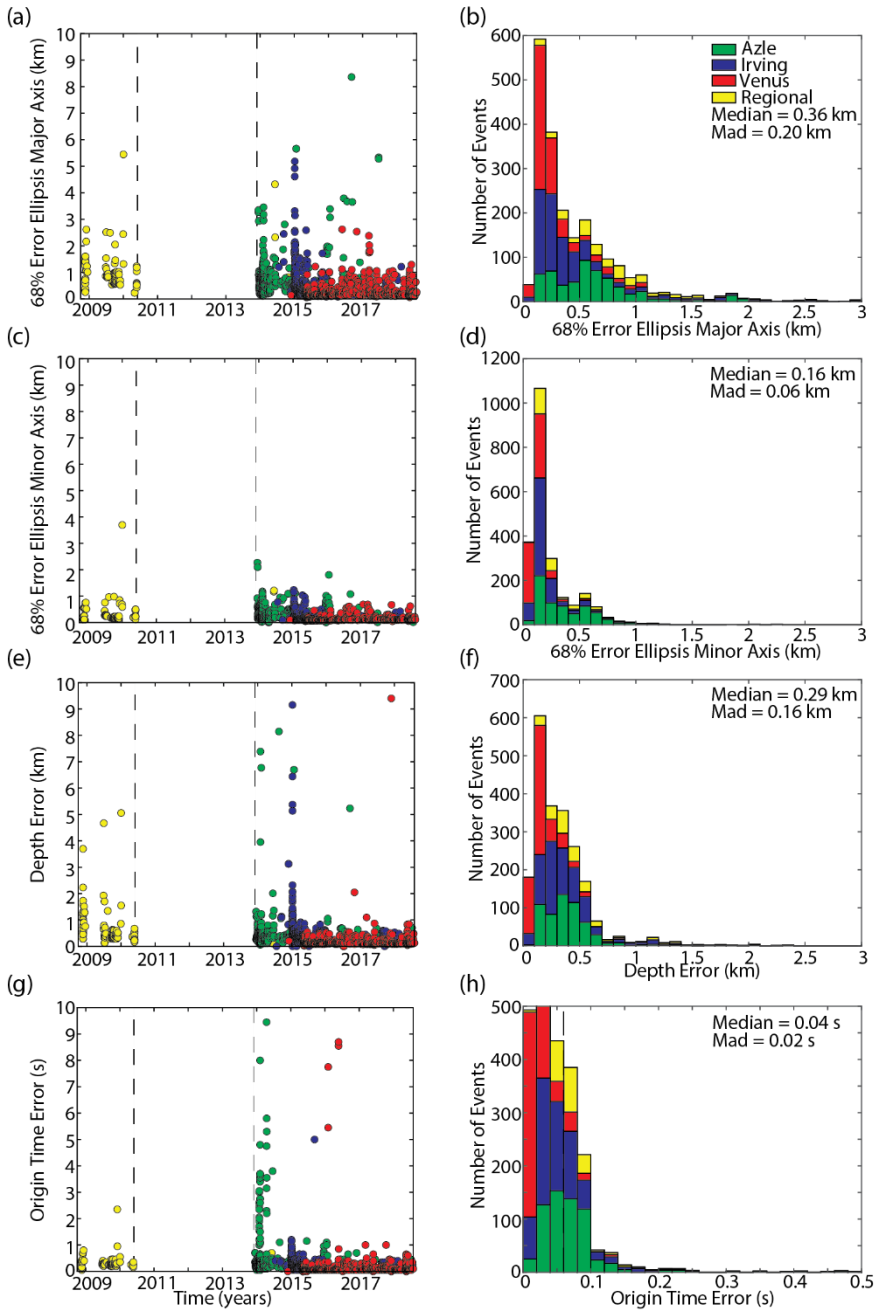

Figure 2.7. (a,c,e,g) Scatter plot views showing the lengths of the major (a) and minor (c) axes of the 68% confidence error ellipsis, associated depth errors (e), and origin time errors (g) of the NTXES catalog earthquakes vs. time (dashed lines indicate time period where there were no active stations within the Fort Worth Basin). Each earthquake is represented by a separate circle shaded by their sequence. (b,d,f,h) Stacked histograms showing the distributions of the same four error parameters in the same order for the NTXES catalog. Each sequence's contribution to the cumulative distribution of location errors are shown. The median and median absolute deviation (MAD) values for each location parameter are also shown.

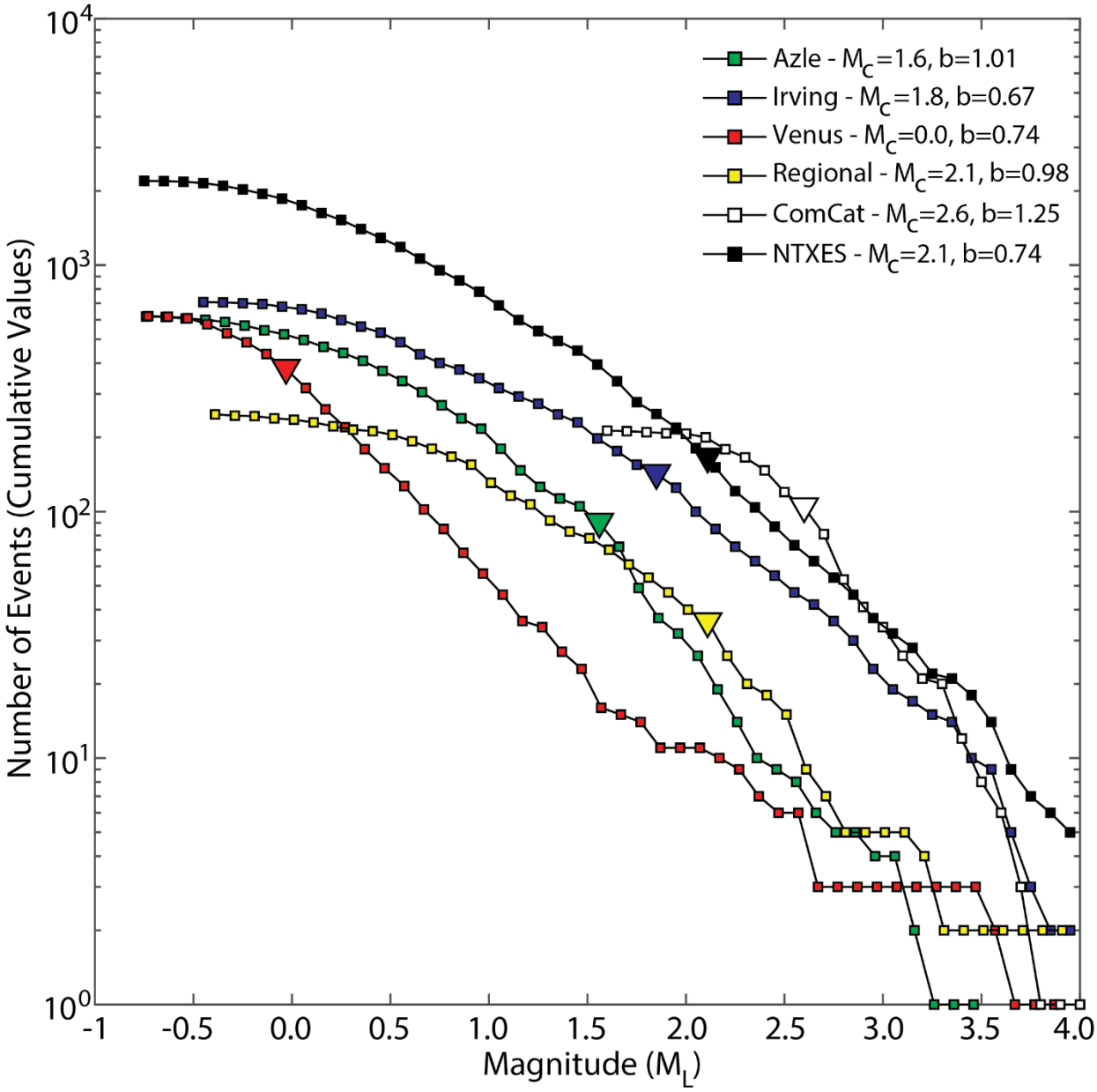

Figure 2.8. Gutenberg-Richter plot showing the magnitude of completeness (Mc) and b-values for each subsection of the NTXES catalog, the NTXES catalog as a whole, and the ComCat catalog of Fort Worth Basin events for comparison. The Mc values are represented by the inverted triangles in the plot.

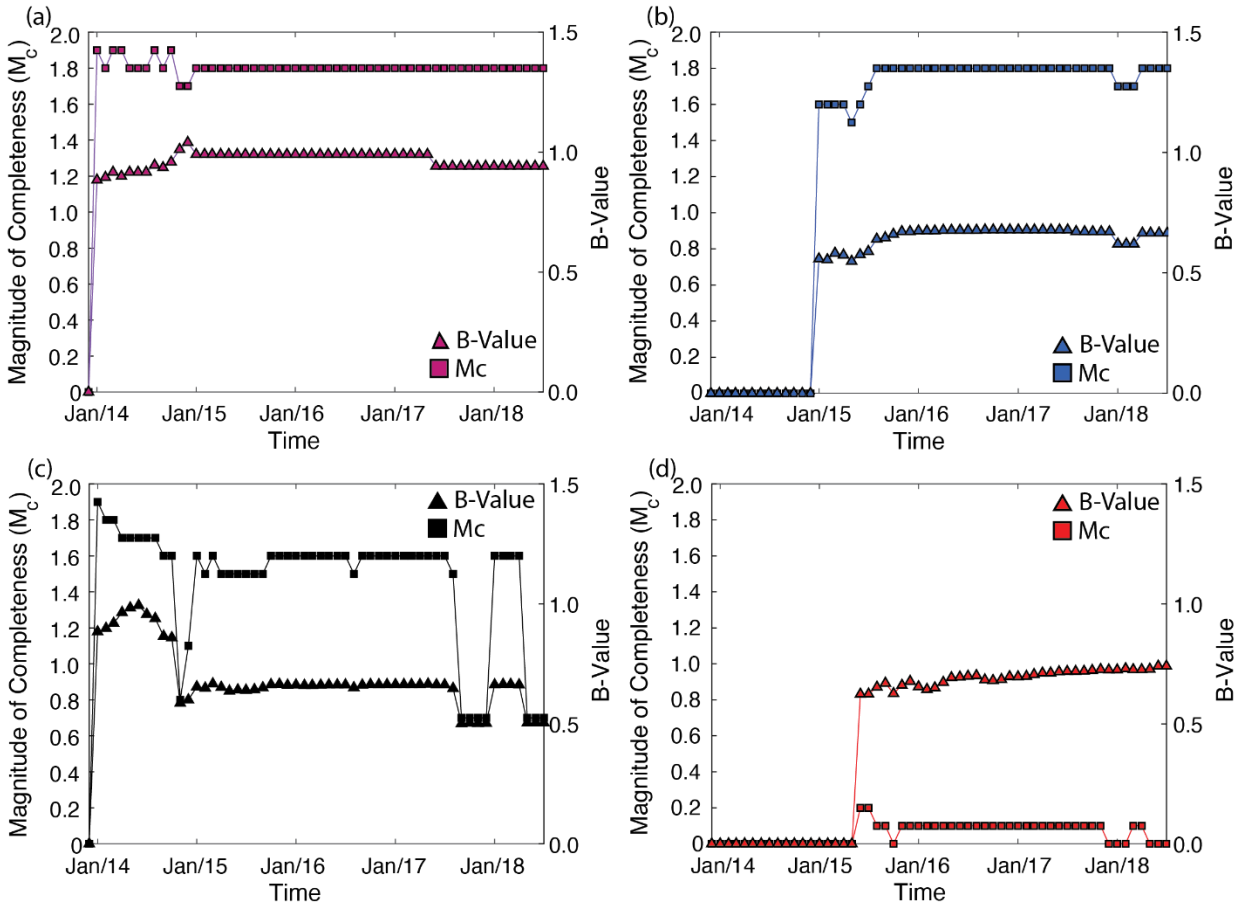

Figure 2.9. Plots showing the temporal evolution of the calculated b-values (triangles) and magnitude of completeness values (squares) where only the events in each subset of the catalog that occurred during and before that temporal bin (bin size = 1 month) for the (a) Azle-Reno, (b) Dallas-Irving, (c) Regional, and (d) Venus subsets of the NTXES catalog.

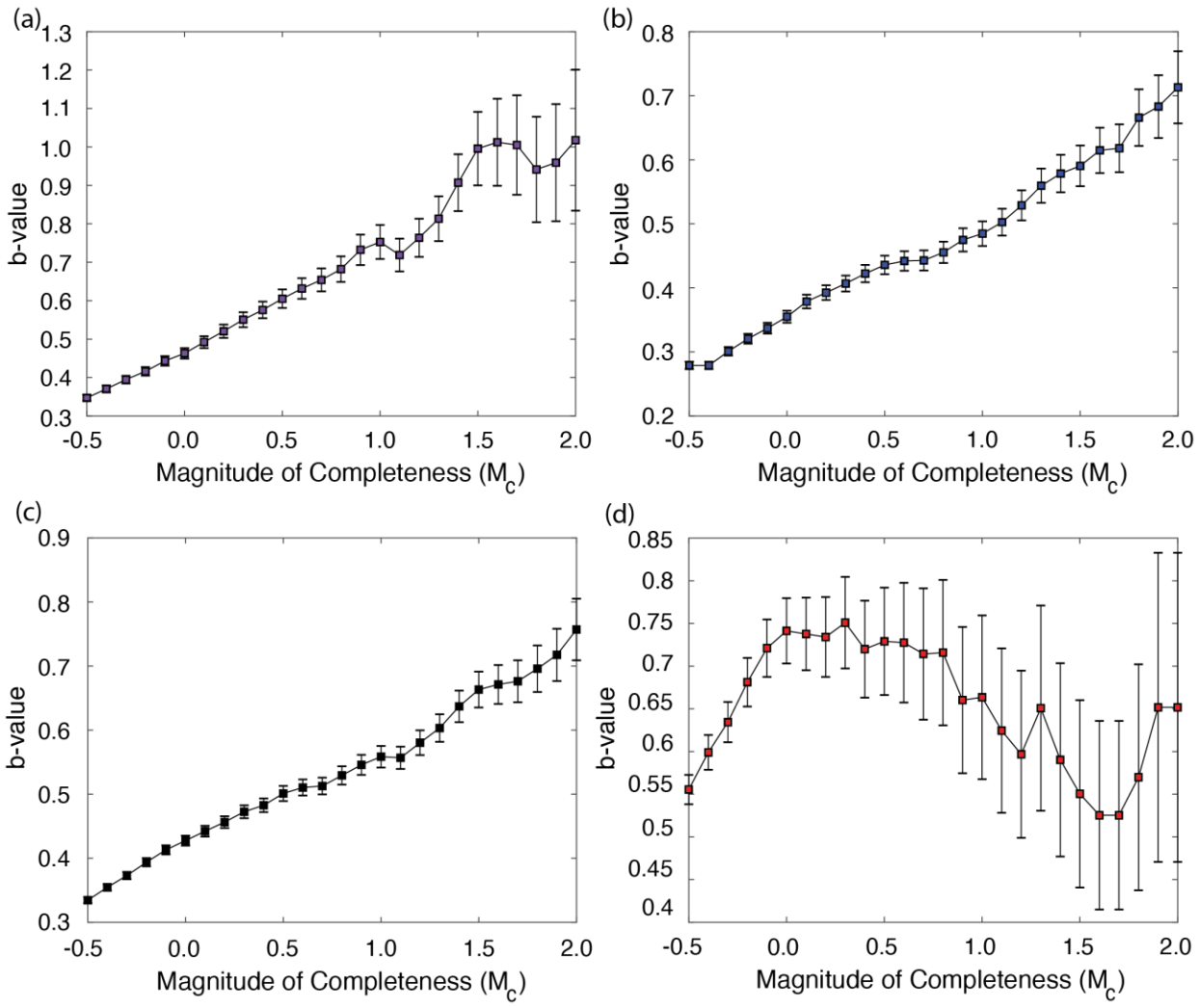

Figure 2.10. Plots showing the calculated b-values for subsets of the NTXES catalog using varying set magnitude of completeness values for the (a) Azle-Reno, (b) Dallas-Irving, (c) Regional, and (d) Venus NTXES catalog subsets.

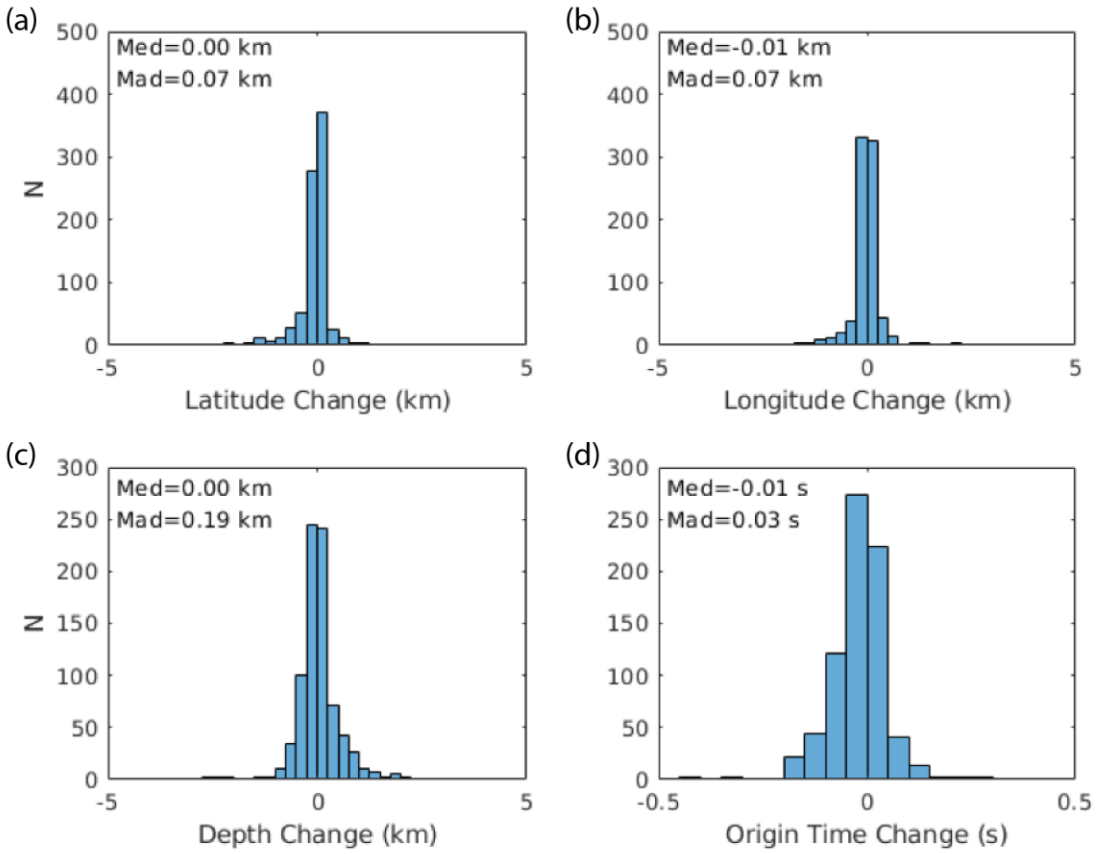

Figure 2.11. Histogram plots showing the distributions of the changes in latitude (a), longitude (b), depth (c), and origin time (d) from the previously published versions of the 2008-2009 DFW airport (Frohlich et al., 2011), 2009-2010 Cleburne (Justinic et al., 2013), 2013-2014 Azle (Hornbach et al., 2015), and 2015-2017 Venus (Scales et al., 2017) to the newly presented NTXES catalog locations. The median and median absolute deviation (MAD) values of each change are also shown. Negative values represent a southward movement in terms of latitude change, a westward movement in terms of longitude change, an upward movement in terms of depth change, and a move to an earlier time in terms of origin time change with positive changes representing movement in the opposite direction for each of the parameters.

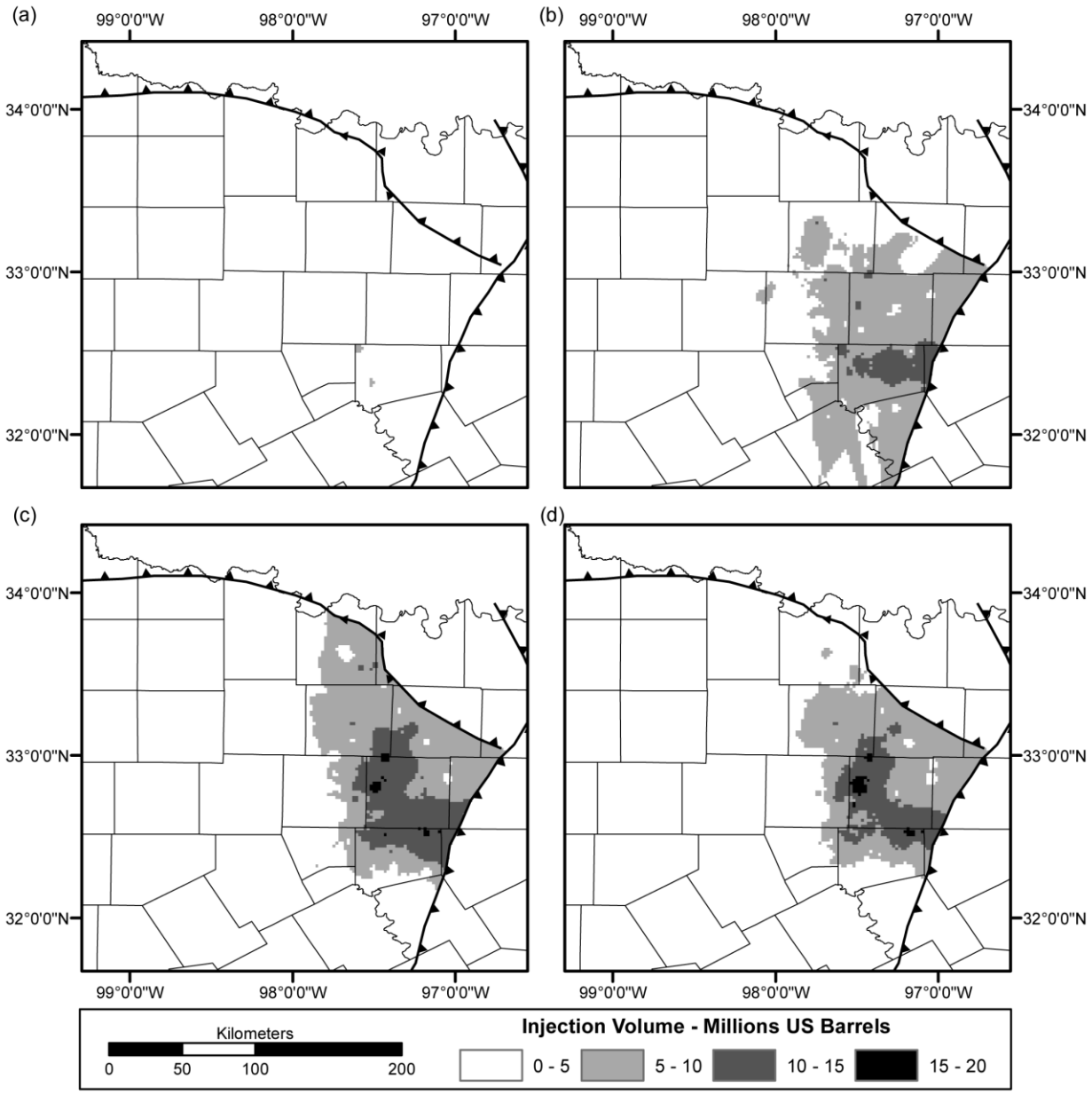

Figure 2.12. Map views of the volumes of injected fluids in the Fort Worth Basin over the time periods of 2005-2007 (a), 2008-2010 (b), 2011-2013 (c), and 2014-2017 (d) using the same interpolation method as described in Figure 2.13.

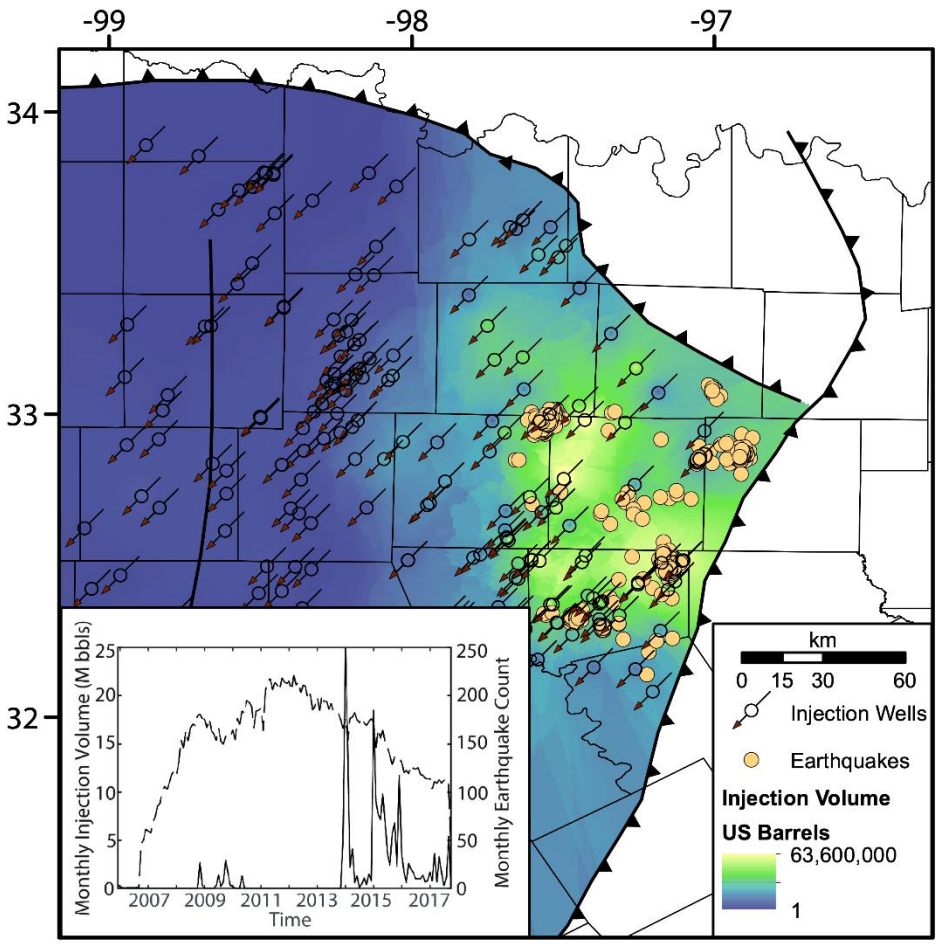

Figure 2.13. Map view showing the interpolated cumulative injection volumes of all fluids injected into the Ellenburger formation from Oct. 2005-Oct. 2017. The interpolation was conducted using an inverse distance weighting scheme using a weighting power of 1 and using data values taken from the 10 nearest wells to each point in space. Each cell is approximately 1.94 km by 1.94 km in size. Also shown are the earthquake (circles) and injection well (arrows) locations. (inset) Plot showing the monthly injection volumes in millions of US barrels (M bbls) for the FWB (dashed line) and the monthly number of earthquakes recorded within the NTXES catalog (solid line) over the same period.

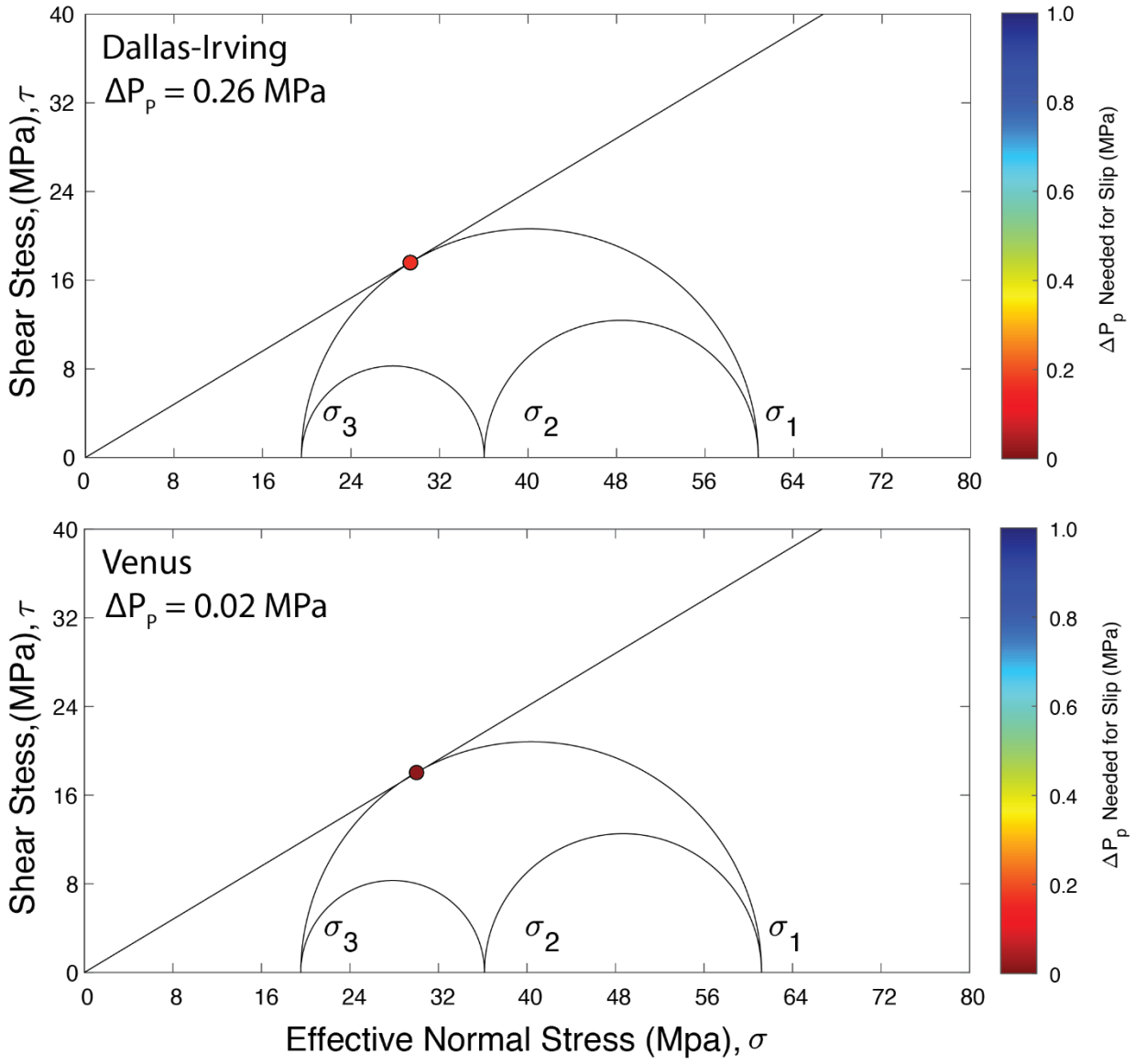

Figure 2.14. Mohr circle representations of the effective stresses per km for the (top) Irving–Dallas and (bottom) Venus sequences. The solid line represents the original failure criterion created using values for the friction coefficient μ of 0.6 and the cohesion C of 0. The circles represent the source fault orientation interpretations created from the combined earthquake and seismic reflection data available at each sequence site. The listed stress changes per km needed for those fault planes to slip and induced seismicity are listed. The depths of each of the earthquake sequences can be interpreted to be 4 km.

3. THE DALLAS-IRVING EARTHQUAKE SEQUENCE

3.1. The Dallas-Irving Earthquake Sequence

Earthquake sequences such as the Azle-Reno and Venus sequences within the Fort Worth Basin are examples of induced seismicity through the analysis of the spatiotemporal correlation between fluid injection activities and seismic activity at each of the fault locations (Frohlich et al., 2010, 2011; Frohlich, 2012; Reiter et al., 2012; Justinic et al., 2013; Hornbach et al., 2015; Scales et al., 2017; Ogwari et al., 2018; Quinones et al., 2019). I asserted that the primary cause of the induced seismicity on most faults in the FWB such as the Venus sequence is direct pore fluid pressure diffusion produced by the fluid injection activities near each of the faults (Scales et al., 2017; Quinones et al., 2018, 2019). However, the earthquake sequences occurring in the most northeastern portion of the basin have remained exceptions to this causal pattern of seismicity within the FWB. The Dallas-Irving earthquake sequence has no fluid injection wells nearby (<15 km distance) the source fault and so lacks a simple linkage to direct pore fluid pressure effects (Figure 3.1). Therefore, I proposed in the previous chapter that analysis of basin-wide stress change effects induced by fluid injection activities would be required to properly assess the main cause of stress change inducing slip on the causative fault of the Dallas-Irving sequence. Here, I present the detailed history of the Dallas-Irving earthquake sequence and the different methodologies employed to better understand both how the basin-wide injections and associated

stress changes have affected the Dallas-Irving fault and how the Dallas-Irving earthquake sequence potentially began.

Due to the lack of fluid injection activities nearby the Dallas-Irving sequence one might consider the idea that the Dallas-Irving earthquake sequence may not be an example of induced seismicity. However, if we examine the Dallas-Irving sequence in the context of the screening methodology proposed by Davis and Frohlich (1993) then we can conclude based on their seven questions that the Dallas-Irving sequence is likely induced. The questions are:

1. Are the events the first known earthquakes of this character in the region?
2. Is there a clear (temporal) correlation between injection and seismicity?
3. Are epicenters near wells (within 5 km)?
4. Do some earthquakes occur at or near injection depths?
5. If not, are there known geologic features that may channel flow to the sites of earthquakes?
6. Are changes in well pressures at well bottoms sufficient to encourage seismicity?
7. Are changes in fluid pressure at hypocentral locations sufficient to encourage seismicity?

The answers to questions 1, 4 and 5 are yes since the Dallas-Irving earthquake are the first recorded events in modern history (since the establishment of the city of Dallas) to occur along this fault (Frohlich et al., 2010; Hornbach et al., 2015; Magnani et al., 2017; Ogwari et al., 2018), the earthquakes in the sequence occur within the crystalline basement beneath the main fluid disposal formation in the basin (Ellenburger formation) (Figure 3.2) (Quinones et al., 2019), and the Dallas-Irving fault creates hydraulic connectivity between the Ellenburger and crystalline basement formations (Hennings et al., 2019; Horne et al., 2020). The answer to question 2 is a

more ambiguous for the Dallas-Irving sequence for while there is no nearby injection well, the Dallas-Irving sequence began following periods of high basin-wide injection activities (Figure 3.3). The answer for question 4 is “no” as again there is no nearby injection well to the Dallas-Irving fault. This mixture of yes and no responses leads to an ambiguous interpretation of the nature of the sequence and requires the quantified assessment of questions 6 and 7 to fully interpret the nature of the Dallas-Irving sequence. In the state of Texas, the primary regulatory agency monitoring oil and gas activities is the Texas Railroad Commission (TRRC) and while this agency does require annual reporting of monthly injection volumes from each active fluid disposal well in the state, the agency does not require the public release of any well pressure information (webapps.rrc.texas.gov/H10/h10PublicMain.do). This means that I do not have access to these parameters even in locations where there is both high volume injection activities and recorded seismicity.

One of the few locations where some publicly available injection pressure data is to the south, near the Venus earthquake sequence to south in Johnson County, where fall-off test data was publicly reported following the occurrence of a M_W4.0 event nearby certain injection wells (Scales et al., 2017). This fall-off test data reported an over-pressured Ellenburger formation with pressure values ranging from 1.7-4.5 MPa above hydrostatic values. No such well pressure data has been reported for the area surrounding the Dallas-Irving sequence and so question 6 cannot be answered using data reported by oil and gas companies. Likewise, exact changes in fluid pressure at the Dallas-Irving sequence site do not exist. Instead, to explore these questions a coupled geomechanical model designed to model both the pore fluid pressure and poroelastic stress changes observed across the FWB, including at the site of the Dallas-Irving sequence, associated with known fluid injection activities was used.

The coupled geomechanical model was created using the COMSOL modeling software and uses the most up to date data sets describing the geologic and hydrologic parameters of the main formations of interest in the basin along with the injection information provided by the TRRC (Inc., C. 2010). The focus of the model is to analyze the cumulative effects of fluid injection activities across a 100 km^2 portion of the FWB where both oil and gas and seismic activities are concentrated (Figure 3.4). Particular attention will be paid to the far field injection related stress changes occurring across the basin as a means of estimating the effect that basin-wide injection activities have had on the Dallas-Irving earthquake sequence. Due to limitations of the input data sets and the modeling software itself, multiple model realizations have been run focusing on the effects of parameters such as the bulk permeabilities of certain key formations of the distribution of permeable faults across the basin on the model results. The modeling results are also compared to the seismic history of the Dallas-Irving sequence by analyzing when the model results predict injection related stress changes reach the source fault of the sequence and when earthquakes associated with the fault were first recorded. However, this temporal analysis of the relationship between basin-wide injection and seismic activities is incomplete as local monitoring of the Dallas-Irving earthquake sequence only began sometime after the onset of seismicity on the fault (Figure 3.5).

The Dallas-Irving earthquake catalog discussed in chapter 2 as part of the North Texas Earthquake Study (NTXES) catalog was created using the data collected from the local monitoring network comprised of stations operated by Southern Methodist University (SMU) and the Texas Seismic Monitoring Network (TexNet) (DeShon et al., 2019; Quinones et al., 2019). As such the earliest associated earthquakes within the Dallas-Irving earthquake catalog coincide with the deployment of local stations by SMU, just prior to the onset of a pair of

M_L3.5+ events associated with the Dallas-Irving fault in January of 2015 but may not have captured earlier seismicity associated with the fault (Figure 3.5). Therefore, the exact start date of the Dallas-Irving earthquake sequence remains in question. To better estimate the start time of the Dallas-Irving sequence, a cross-correlation based template matching approach using the *EQcorrscan* code package was applied to all seismic data recorded in the FWB with stations deployed before January 2015 to detect any missed events prior and during the monitoring period for the Dallas-Irving sequence (Chamberlain et al., 2017). This process involved analyzing both station waveform data from the local monitoring stations and from stations operated at regional distances from the Dallas-Irving fault which were active prior to January 2015 (Figure 3.6). After creating an updated list of detected events using the template matching based approach, the updated set of event detections were used to interpret the modeling results to gain a better understanding of the temporal relationship between the fluid injection and seismic activities occurring at the Dallas-Irving sequence site.

3.2. Geology of the Dallas-Irving Site

The FWB foreland basin formed approximately 300 Myr ago during the Ouachita orogeny and is bounded to the south by the Llano uplift feature, to the west by a general shallowing of formations, to the north by the Red River and Muenster arch features, and to the east by the Ouachita Thrust Front feature (Pollastro et al., 2007). For the purposes of analyzing the Dallas-Irving sequence this eastern boundary is the most important to consider because of its close location to the seismicity and its possible role as a barrier to fluids. The Ouachita Thrust Front is described as a shallow ESE-dipping fault overlain by denser metasedimentary units on top of other major formations which characterize the eastern margin of the FWB basin including the Dallas-Irving sequence site (Figure 3.7) (Pollastro et al., 2007; Magnani et al., 2017;

Hennings et al., 2019). Additionally, the Ouachita Thrust Front is considered the eastern boundary for hydraulic connectivity of the Ellenburger formation that in modeling terms can be treated as an impermeable barrier to fluid flow and impact estimates of associated stress changes. Most production related interpretations of the basin place the barrier to fluid flow within the Ellenburger formation at the location of the interpreted contact between the bottom of the Ellenburger formation and the Ouachita Thrust Front (Zhai and Shirzaei, 2018; Chen et al., 2020, Haddad and Eichhubl, 2020; Gao et al., 2021). Other geologic interpretations of the hydraulic connectivity of the Ellenburger formation across the Ouachita Thrust Front exist, including one interpretation near the Venus sequence that asserts there is limited hydraulic connectivity for upwards of 10's of km past the mapped Ouachita-Ellenburger contact point (Figure 3.8; Magnani et al., 2017). Unfortunately, a geologic interpretation made using well log data spanning 10s of km past the mapped Ouachita-Ellenburger contact point does not exist for the portion of the FWB closer to the Dallas-Irving sequence or for other areas of the basin between the two earthquake sequences. As such, there is some uncertainty as to how this eastern boundary should be treated within the model with regards as to its spatial location and whether the boundary should be treated as an entirely impermeable feature. While both endmember constructions are tested, in the end, I decided to interpret the production focused construction that treats the mapped Ouachita-Ellenburger contact point as the eastern boundary of the basin and as an impermeable feature. This interpretation better aligns with interpretations of the geology of the Fort Worth Basin (Smye et al., 2018) and serves to allow for better comparisons between the modeling results presented here and those presented in other publications (Zhai and Shirzaei, 2018; Gao et al., 2021). Additionally, differences in certain hydrologic parameters and the

inclusion of poroelasticity can be assessed in terms of the resulting stress changes (Zhai and Shirzaei, 2018; Gao et al., 2021).

3.3. The Dallas-Irving Earthquake Catalog

The first events believed to be associated with the Dallas-Irving earthquake sequence were recorded using regional stations operated by SMU within the FWB and by the USGS using their permanent array (Figure 3.5). There exists some uncertainty about whether these events which occurred in late 2014, were occurring along what is now referred to as the Dallas-Irving fault due to larger location uncertainties owing due to the lack of nearby stations. A complicating matter is that seismicity in October 2008 had previously been recorded on a separate fault to the west beneath the Dallas-Fort Worth International Airport (Frohlich et al., 2011; Ogwari et al., 2018). Initially, it was believed that these later 2014 events could have been occurring along the Airport fault, but analysis of the S-minus-P times, particularly at station DAL at SMU, revealed distance discrepancies that placed the events further to the east and closer to SMU (C. Hayward and H. DeShon, personal communication). Following this observation, on January 5 SMU deployed a broadband sensor at the DFW Airport and a day later the pair of M_L3.5+ events on the Dallas-Irving fault occurred (Figure 3.9). During the subsequent weeks, the number of stations was rapidly increased; first by quickly installing a set of one-component nodal stations in the area and by late January 2015 installing additional three component sensors borrowed from the Incorporated Research Institutions for Seismology (IRIS) and the USGS that included short-period, broadband, and strong motion NetQuakes instruments. Since then, the local network monitoring the Dallas-Irving earthquake sequence has changed dramatically with sensors being relocated, installed, and eventually, following a steady decrease in the seismicity rate of the fault, removed (see DeShon et al. 2018). By 2021, at the time of this writing, SMU operates 5 stations

over this sequence. Another major change in the monitoring of the Dallas-Irving sequence came in 2016 with the deployment of the regional TexNet stations designed to improve monitoring across the entire FWB (Savvaidis et al., 2019). These broadband sensors were installed across the entire extent of the seismically active portion of the basin with the stations closest to the Dallas-Irving sequence incorporated into the monitoring efforts of the named sequence by SMU.

The SMU Dallas-Irving earthquake catalog was created using the Antelope environmental monitoring system (Boulder Real Time Technologies, brtt.com) and is comprised of those events associated with the singular active Dallas-Irving fault. The catalog spans a period from 2014-2021 for which SMU operated stations monitored the sequence. In total there are 750 events associated with the Dallas-Irving sequence with magnitude ranging from M_L-0.5 to M_L3.6 (Figure 3.5). The events are all associated with a singular NE-SW trending steeply dipping normal fault with a seismically active area of approximately 12 km^2 (Figure 3.2). The earthquakes are primarily located within the crystalline basement formation that underlies the primary fluid injection formation in the basin, the Ellenburger limestone-dolomite formation (Quinones et al., 2019). Most seismicity associated with the Dallas-Irving fault occurred during the first year of activity and displayed a circular pattern of earthquake migration around the larger early earthquakes, indicating that different parts of the fault appeared to become active at different points in time (Figure 3.10). The overall trends in seismicity were that: the initial events occurred along the southern portion of the fault, aftershocks migrated north and depths of the initial events were generally shallower than those of the later events. Focal mechanism solutions created for the Dallas-Irving sequence data show almost entirely normal faulting solutions with an average strike/dip/rake orientation of 39°/69°/-88° (Figure 3.11) (Quinones et al., 2018). After performing a formal stress inversion of the focal mechanism catalog collected from the eastern

margin of the basin from the Dallas-Irving and Venus sequences we found that this portion of the basin was operating under a normal faulting stress regime with a maximum horizontal stress orientation of 25° E of N (Quinones et al., 2018). This means that the Dallas-Irving fault was critically stressed within the currently active local stress field and thus would not have required a large stress change to initiate slip on this fault associated with the earthquakes. As presented in Chapter 2 (this dissertation), inversion of the focal mechanisms in fact suggests as stress changes as low as 0.26 MPa would be sufficient.

3.4. Double-Difference Relocation

The double-difference earthquake relocation technique is applied to this data to gain a better understanding of the distribution of seismicity and the geometry of the Dallas-Irving fault. The specific double-difference location algorithm used was the *tomoDD* set of codes which combines cross-correlation differential times, catalog differential times, and absolute catalog travel times to better constrain the absolute and relative earthquake locations (Zhang and Thurber, 2003). The basis of the double-difference relocation methods relies on the use of differences in travel times, known as differential times, to calculate relative earthquake locations under the assumption that waveforms sourced from earthquakes from similar locations follow similar ray paths and that differences in travel times across recorded earthquakes reflect differences in earthquake locations (Waldhauser and Ellsworth, 2000; Zhang and Thurber, 2003). One of the key benefits to *tomoDD* is that the location algorithms do not remove earthquakes that are not well linked to other events in the catalog and does not require new clusters to be defined for those unique events.

The methodology employed to run *tomoDD* follows a similar procedure outlined in Scales et al. (2017) and described in detail here. The first analysis step following the creation of the Dallas-Irving earthquake catalog was to calculate the absolute travel times for each event to each detecting station and the differential times for the entire catalog. The program *ph2dt* was used in making these calculations (Waldhauser and Ellsworth, 2003). In order for a set of catalog differential times and absolute travel times from a specific event to be included in the double-difference location algorithm, the following parameters were used: an event needs a minimum of at least 8 phase arrivals (either P- or S-phase arrivals); an event pair needs to be located within 5 km of at least four other event pairs; and the station to event distance must be less than 30 km. An azimuthal gap coverage parameter was also implemented where events with azimuthal P-wave separation values of >180° were removed (e.g., Koulakov, 2009). The implementation of these parameters decreased the total number of events in the Dallas-Irving catalog to 600 with mostly the smallest magnitude events which lacked a greater number of phase picks being removed from the initial catalog.

Cross-correlation differential times for the Dallas-Irving earthquake sequence data were calculated using the University of Alaska Geophysical Institute's Seismology Matlab Objects Suite for MATLAB known as GISMO (Reyes and West, 2011). GISMO allows for the transfer of the database produced in Antelope into a MATLAB environment for subsequent processing. The cross-correlation values and differential times were calculated using different parameters for the P- and S-wave based on the different frequency bands where the signals are strongest with the P-wave energy being slightly more impulsive than the S-wave energy across the Dallas-Irving network. Signals were processed using a bandpass Butterworth filter from 10 to 40 Hz with a time cut from 0.25 s before to 0.75 s after the P-wave arrival time and a 2 to 25 Hz filter

with a time cut from 0.5 s before to 1.0 s after the S-wave arrival time. Cross-correlation calculations are computed for every station channel that recorded at least 10 phase arrivals. The calculation of the cross-correlation differential times followed several steps. First, all arrival times associated with a station-channel pair using the defined time cut windows for the corresponding phase (Figure 3.12) were extracted followed by the application of the filters. Next, the data was transformed into the frequency domain and each individual phase arrival from the station-channel pair was cross-correlated with all other phase arrivals recorded at the station yielding a cross-correlation matrix plot (Figure 3.13). Lastly, the data was transformed back into the time domain and the lag times associated with the maximum cross-correlation values from each event pair were used to define a lag time value used to adjust the waveforms of one of the two arrival times from each event pair of catalog differential times (Figure 3.14-3.16). After the waveforms were adjusted, the differential times are calculated and reported as the cross-correlation differential times. Additionally, clustering analysis of the events recorded on each station based on the cross-correlation values of each event to another was performed (Figure 3.17). Based on the clustering analysis, the cross-correlation differential time list was refined to include only those event pairs with a minimum cross-correlation coefficient of 0.6 with at least 8 phase arrivals associated with the event pair. Inspection of the cross-correlation differential times and the adjusted waveforms identified evidence of cycle skipping, which occurs when the phase match between two waveforms is greater than half a wavelength, associated with certain events (Figure 3.16). To assess the impact of cycle skipping on the cross-correlation differential time, these results were compared to their associated catalog differential times for each event pair (Figure 3.18). This comparison showed that there were several event pairs whose cross-correlation and catalog differential times significantly differed from one another that could

negatively impact the *tomoDD* double-difference location algorithm. As a result of this analysis, cross-correlation differential times from these event pairs which had a greater than 0.2 s difference in values between the cross-correlation and catalog differential times were eliminated from further use.

I utilized *tomoDD*'s least squares methodology for solving a linearized system with QR factorization to determine the relative earthquake locations (Paige and Saunders, 1982). Here, *tomoDD* was run for a total of 12 iterations using five different weighting schemes with a variable number of iterations associated with each weighting scheme (Table 3.1). The weighting scheme heavily favors the absolute travel times in the first set of iterations (1-2), favors the catalog differential time data in the second set of iterations (3-4), upweights both the absolute travel time and catalog differential time data in the third set of iterations (5-6), before upweighting the cross-correlation differential time data in the final two sets (7-9 and 10-13). The resulting output of the *tomoDD* analysis is shown in Table 3.2. I observed large decreases in the overall root-mean-square (RMS) residual values of both the catalog (RMSCT) and cross-correlation (RMSCC) data within each set of iteration steps and observe that the average location of the events does not shift by more than 10s of meters after only three iterations (DX, DY, and DZ). When examining the locations of the relocated events over the same period that the Dallas-Irving sequence was most active (2015), I observe that the initial event locations which were originally much shallower, but with higher location uncertainties, have moved closer to the main cluster of seismicity (Figure 3.19). Aside from the earliest event locations from within the NTXES catalog, which had the highest location uncertainties in that named catalog, the rest of the 2015 events have become more tightly clustered when compared to their original NTXES catalog locations shown in Figure 3.10. The migratory pattern for different parts of the fault

becoming active at different points in time noted in the original NTXES catalog remains (Figure 3.19). Overall, the double-difference relocated earthquake catalog is very similar to the initial Dallas-Irving catalog, however many of the outlier events and the main cluster of seismicity have become more tightly clustered along the assumed fault structure (Figure 3.20). The Dallas-Irving double-difference earthquake catalog information is presented in Appendix B.1.

3.5. EQcorrscan Template Matching

Template matching is an event detection technique where typically an analyst uses a detected event signal captured on a particular sensor to scan through segments of waveform data collected by that same sensor to search for similar waveform patterns in order to detect similar events, possibly smaller in size (Gibbons and Ringdal, 2006; Shelly et al., 2007; Peng and Zhao, 2009; Skoumal et al., 2014, 2020; Ross et al., 2017; Beaucé et al., 2018; Chamberlain et al., 2018). The template matching technique implemented here is *EQcorrscan*, which is cross-correlation based event detection procedure found in a set of codes built using Python (Chamberlain et al., 2018). The full workflow of the codes is described by Chamberlain et al. (2018) and summarized here. First, a set of "template" events were defined. These templates were built using the associated phase detections compiled from an already detected and located event within an earthquake catalog. Here, the templates used in *EQcorrscan* were based on events from the Dallas-Irving portion of the NTXES catalog which have not only event origin time and location but also the individual station phase detection times stored as in Antelope (Figure 3.21). The relevant event and phase detection information were extracted from the Antelope database and converted into a full QuakeML catalog that could be loaded into the Python *EQcorrscan* module. The waveform data were accessed using the IRIS webservices tools built into the *obspy* Python module and were processed using the same frequency filters as were used to detect the original events

(Figure 3.22; https://service.iris.edu/). Once the full set of templates was constructed from the Dallas-Irving event catalog they were combined into a "Tribe" object which is simply a collection of template events. Once the templates were constructed the waveform data for the time periods to scan for events was defined and accessed, again using the IRIS webservice tools and processed the waveform data using the same parameters as those defined for the templates. The result of this template matching analysis was a list of event detection times with an overall cross-correlation value associated with the template in the Tribe that the event detection is most strongly correlated (Figure 3.23). Additionally, the sensors from the template that detected the event were tracked along with their respective cross-correlation values from the specific phase detection. These cross-correlation values were next used in cross-correlation based phase picking on these sensors to determine actual phase arrival times using the lag calculation function provided within *EQcorrscan*. The resulting product is a set of events with overall event detection times and specific phase arrival times for each sensor with a high enough cross-correlation value to a phase picked on that same sensor from the template event for each event.

The goal of using the *EQcorrscan* codes to examine the Dallas-Irving earthquake sequence data is to gain better insight into when the Dallas-Irving sequence began. *EQcorrscan* was run across the full operating history of stations active within the FWB from late 2013 to the present. The analysis provides a list of event detection times and cross-correlation based phase picks for each of the events that can be used to make earthquake location estimates for the template-matching detected events. A future goal of performing double difference earthquake relocation using these template matched events and the NTXES earthquake catalog may be possible if phase arrival times are included.

Due to the complexity in the monitoring history of the Dallas-Irving earthquake sequence, there were needed modifications to the implementation of the *EQcorrscan* codes. To define the parameters used to create the templates, I needed to compare the Dallas-Irving dataset to the data sets discussed in other publications using the *EQcorrscan* codes. The main comparison data set was the work completed using the real-time version of the *EQcorrscan* codes wherein the template matching technique was used to examine events off the coast of New Zealand (Chamberlain et al., 2020). The major difference between the two data sets was that the Dallas-Irving sequence has had numerous monitoring stations become active for only small periods of time with numerous sensors moved and renamed over time with the complex history of the monitoring efforts for the Dallas-Irving sequence presented in DeShon et al. (2018). Thus, it was important to know which stations were active and recording events when deciding which templates are to be used with what period of waveform data (Figure 3.9). My methodology to account for these effects was to separate the Dallas-Irving sequence dataset into distinct time periods with their own associated sets of templates reflecting which stations were active for each period. Additionally, the Dallas-Irving earthquake catalog is mainly composed of small magnitude ($<M_L1.0$) events which are tightly clustered together (Figure 3.1). Therefore, the events from the Dallas-Irving catalog that were transformed into templates used a magnitude ($>M_L1.0$) and distance (<1 km) cutoff to accompany the temporal separation of the catalog. These criteria led to 51 templates that covered five distinct periods: the pre-2015 period where there were no local monitoring stations with events detected and located primarily using the Azle-Reno stations; the period covering the first 10 days of 2015 where the number of broadband and short-period stations were greatly limited and one-component nodal stations were active; the remainder of 2015 where monitoring was dominated by a local network composed of a mixture

of short-period, broadband, and NetQuakes strong-motion stations; the period covering the first half of 2016 when the NetQuakes were demobilized but the short-period and broadband stations remained active; and the final period from mid-2016 to the present day where the monitoring of the Dallas-Irving sequence included the broadband stations installed by TexNet (Figure 3.24). Another major difference from the Chamberlain et al. (2020) studies was the definition of the radius around a possible event detection from which to pull templates from. In the New Zealand study when they pull an event detection with an associated location into their workflow, they proceeded to utilize all created templates from events within at least a 50 km radius to the detected event epicenter to search for other possible events. This template search radius was far larger than what was required in the FWB, where using a 50 km radius around a possible Dallas-Irving event would mean templates associated with the Venus earthquake sequence would be included in the search for missing events. While I have not examined the cross-sequence cross-correlation values from stations that have detected events on both the Dallas-Irving and Venus faults, the fact remains that both faults are >40 km apart from one another and currently there exists no direct evidence of events occurring along one fault potentially triggering events on the other (Quinones et al., 2019). As such, the search radius around the Dallas-Irving sequence site from which I pull template events was reduced to 15 km.

Following the creation of the templates from the selected Dallas-Irving events, the next step was to access the continuous waveform data collected for the stations in the FWB with associated template detections. A problem arose at this step using *EQcorrscan* codes due to unique ID value assigned to each station-channel pair used within the templates. The standard methodology for ensuring a specific sensor has a unique ID value at IRIS is to assign a station name, channel name, network name, and location code to each sensor. In the cases for the FWB

where a station needed to be moved or a sensor was replaced the changes were accounted for within the Antelope database and at IRIS by either changing the station name or the location code for the instrument. However, the standard methodology for assigning ID values for a sensor in the *EQcorrscan* codes was to only examine the station and channel names. Therefore, I needed to manually add a function to gather the waveform data associated with the specific stations and channels used to create the templates for each period which included additional information about the network and location codes to properly access the data at IRIS. Once this change was implemented into the workflow, the five time periods were used to separate out templates for each of the time periods where the associated stations and channels used to create each template were active (Figure 3.24). This portion of the workflow was completed using the high-performance computing center ManeFrame II at SMU. The continuous waveform data collected on the stations with associated templates was analyzed on an hour-by-hour basis using 24 individual 250 GB memory 36-core nodes. On each node, an hour of data was analyzed in approximately 50 seconds, and the entirety of the FWB station waveform data from late 2013 to the present day was analyzed in under 48 hours using the 24 available nodes. The result of this processing step was a list of event detections; each detection having a unique ID value reflecting the detection time of the event and the origin time of the template which most strongly correlated with the event (Figure 3.25). Additionally, each individual station-channel pair which was used to detect the events is also reported along with the cross-correlation value between the detected and template waveforms. Approximately 1700 events were detected over the 6.5-year monitoring period of the Dallas-Irving earthquake sequence which represents an increase of nearly 1000 events compared to the original NTXES Dallas-Irving catalog.

The final processing step taken in *EQcorrscan* was to use the cross-correlation values for each of the station-channel detections to determine their phase arrival times. This process was completed by first setting a minimum cross-correlation value threshold for calculating a phase arrival time. These lag-time corrected phase arrival times are reported along with the refined list of event detections. What is not included with these event detections are origin time or location estimates for these events. In the New Zealand study Chamberlain and coauthors (2018) reported the location of the event detection as that for the best matched template, which is what is done for the Dallas-Irving catalog in Figure 3.26.

It is possible to extract the phase arrival and event detection times out of the *EQcorrscan* codes and potentially use them in a separate earthquake location algorithm in order estimate actual event locations for each of the detections. Many of the event detections, however, lack numerous phase arrival times and so such a procedure would likely lead to poor location estimates. Therefore, I limited the list of event detections to only those events with at least eight phase arrival times which dropped the number of detected events to 820. Ultimately, it seems necessary to include outside phase picking, association, and location algorithms into the workflow to fully convert the template matched earthquake detection list into a full earthquake catalog, and this is left for future work.

3.6. Template Matching Results

The template matching approach provides a better estimate of the initiation of seismicity in the Dallas-Irving sequence site relative to the January 2015 peak of seismic activity (Figure 3.25). There are possible Dallas-Irving earthquakes throughout 2014, almost a full year prior to the installation of the local monitoring network. However, these earliest events most strongly correlated with are templates pulled from the late 2014 and initial weeks of 2015 NTXES

catalog, which by default have the highest location uncertainties in the Dallas-Irving area because of the limited number of regional stations (Figure 3.27). Therefore, the possibility remains that some of these template-matched events detected as early as February of 2014 could have been associated with the DFW airport seismicity, which was still ongoing at the time (Ogwari et al., 2018), rather than the Dallas-Irving fault. Further template matching calculations utilizing the USGS station data would be required to attempt to find possible earlier pre-2014 Dallas-Irving events, however the issue of large location uncertainties associated with the templates used to detect those events would remain and possibly be increased; such a regional template study has been completed by M. Brudzinski and reported in Hennings et al. (2021). Ultimately, I found that the *EQcorrscan* codes did allow me to increase the completeness of the Dallas-Irving catalog and found that the sequence was likely active for months prior to the installation of the local network. However, to truly understand whether these earlier detected events are Dallas-Irving events I would need to incorporate a full suite of outside phase association and event location algorithms into the *EQcorrscan* workflow.

3.7. Coupled Geomechanical Model – COMSOL Foundation

The model presented here is a 3D fully coupled, geomechanical model created using the COMSOL multi-physics modeling software which incorporates both fluid injection associated pore fluid pressure and poroelastic stress changes (Inc. C., 2020). This modeling is accomplished by utilizing the solid mechanics, Darcy's Law fluid flow, and poroelasticity modules within the software. Thus, to explain the construction of the coupled geomechanical, the three named modules incorporated into the software will be discussed.

The Solid Mechanics module is used to define the elastic properties of the various formations incorporated into the model along with the mechanical boundary conditions imposed

upon the model. Within COMSOL all geometric objects in a model are described as "Materials" whose properties along with their size and shape are defined by the user. The Solid Mechanics module requires assignment of Young's modulus, Poisson ratio, and density value for all materials within a model for which elastic stresses will be calculated. The full range of elastic properties for each of the major formations in the FWB used in the model were taken from the collected well log data (Smye et al., 2019) and analyzed samples (Loucks et al., 2009) along with general elastic property ranges compiled from global rock property data sets (Bear, 2013; Wang, 2017) included in Table 3.3. A major focus of the Solid Mechanics module is the calculation of strain and displacements induced within or on a material resulting from an applied load. These calculations are completed using a modified version of Hooke's Law where the Cauchy stress tensor (S) is a function of some load stress (S_{ad}), the stiffness tensor (C) and the elastic strain tensor (ϵ_{el}) (Cosserat and Cosserat, 1909):

$$S = S_{ad} + C : \epsilon_{el} \tag{3.1}$$

Additionally, the strain tensor is defined by the displacement vectors within the model space induced by the injection of some mass, and the stiffness tensor is a function of the elastic properties of each model material. In this model the applied load to the material is the pore fluid pressure changes induced by the injection of a fluid mass into the materials making up the elastic geologic formations. The boundary conditions in the Solid Mechanics module allow the definition of the strain and displacement effects on the edges of the Materials within the model (Figure 3.28). The boundary conditions defined in this model are that the top-most surface of the model along with the boundaries between the different Material layers are free surfaces meaning they have no constraints on the deformation they experience due to loads acting upon those surfaces. The bottom most surface of the model, representing the deepest portions of the pre-

Cambrian crystalline basement rocks is defined as a fixed boundary meaning no displacement or deformation may occur along that surface. The lateral outermost edges of the model are defined to be "rollers" within COMSOL meaning that these surfaces are not allowed to be displaced in the perpendicular direction to the surface, but that the surface is allowed to be sheared in a tangential direction. Lastly, all materials within the model are taken to be isotropic and thus have the same elastic properties in all directions.

The Darcy's Law fluid flow module controls the hydrologic parameters of the fluids and geologic materials within the model and controls the injection of mass portion of the model. In this model, injections of fluids into the Ellenburger formation are treated as a time dependent influx of mass into the elastic material associated with the Ellenburger. The material properties required by the Darcy's Law module are porosity (ϵ_p) and permeability (κ) of the material in question and the density (ρ) and dynamic viscosity (μ) of the fluids flowing through the material. The full equation relating these material and fluid properties to a mass flux is as follows:

$$\frac{\partial}{\partial t}\left(\epsilon_p \rho\right) + \nabla \cdot (\rho u) = Q_m \tag{3.2}$$

Where, **u**, is taken to be a function of the material and fluid properties as well as a function of the pore fluid pressure within the system:

$$u = -\frac{k}{\mu}\nabla p \tag{3.3}$$

The boundary conditions controlled by the Darcy's Law module define whether fluids are allowed to flow across the surfaces on the edges and between each material layers (Figure 3.29). The boundary conditions in this study are defined to allow the free flow of fluids across most surfaces between material layers, while the outermost surfaces of the model space are defined as

no-flow boundary conditions meaning injected fluids cannot exit the model space. An important feature of the model is that while the true model space of interest is only a 100 km^2 area, representing the northeast portion of the FWB where both injection and seismic activities are concentrated, I also include a material layer surrounding the model space which within COMSOL is defined to be a material with infinite lateral length (Figure 3.30). This outermost material layer is included to allow for the far field injection associated stress changes to reduce to zero at large distances from injection wells. The various injection wells, which are taken to be positive mass fluxes as fluids are added to the elastic materials, are modeled as cylindrical objects where the bottom-most surface of the cylinder is taken to be the injection point. The full list of fluid and hydrologic material properties utilized by the Darcy's Law module are included in Table 3.3.

Within COMSOL the Poroelasticity module is a Multiphysics interface meaning that it takes information from other modules within the model and couples them together to solve equations relating to their outputs. Here, I use the Poroelasticity module to couple the Solid Mechanics and Darcy's Law module to solve for the poroelastic stress changes induced by fluid injection activities across the FWB. To begin solving for the poroelastic stress changes I first define a separate set of material properties for each of the fluid and formation materials included in the model (Table 3.3). These properties include the fluid compressibility and the Biot-Willis coefficient value for each of the formation materials. The Biot-Willis coefficient value, defined as being the proportion of fluid pressure that counteracts the confining stress, is taken to be one for all formation materials and the compressibility of the injected fluid is taken to be that of water (Biot and Willis, 1957). Using the compressibility of fluid, the storage coefficient for each formation material in the model is calculated using the equation:

$$S = \epsilon_p \chi_f \tag{3.4}$$

Which can then be related to time dependent changes in pressure within the model system using:

$$\frac{\partial}{\partial t}(\epsilon_p \rho) = \rho S \frac{\partial p}{\partial t} \tag{3.5}$$

The coupling between the Solid Mechanics and Darcy's Law modules occurs at the domain level wherein the calculated pore fluid pressure changes from equation (2) induced by the fluid injection activities are treated as a new applied load upon the elastic formation materials. The full coupling equation utilized is:

$$\rho_f S_p \frac{\partial p_f}{\partial t} + \nabla \cdot (\rho_f u_f) = Q_m - \rho_f \alpha_B \frac{\partial \epsilon_{vol}}{\partial t} \tag{3.6}$$

Where all variables with the *f* subscript refer to the fluid material properties and the volumetric strain (ϵ_{vol}) is calculated within the Solid Mechanics module. Thus, with these three modules the model can calculate the time dependent stress change across the model space induced by fluid injection activities within my specified model geometry.

3.8. Fort Worth Basin Geology and Injection Data

The geologic and hydrologic parameters for each of the major units used in the model are described in the Table 3.3 and are based on the most up to date available data for each of the units. The formation top data for the Ellenburger and Basement formations were collected by colleagues at the Bureau of Economic Geology and is available as part of Smye et al. (2019). The full procedure for processing the digital logs from the nearly 1300 wells active in the FWB are described in Smye et al. (2019) and summarized here. Approximately 1023 of the wells penetrated the Ordovician Viola Simpson and Ellenburger units in the basin while only around 100 of the wells penetrated or nearly penetrated the crystalline basement formation. The

formation tops of each of the major formations of interest were mapped using gamma ray, resistivity, bulk density, neutron porosity and photoelectric factor logs. The formation top elevation values collected at each well site were then interpolated into a surface using the natural neighbor algorithm available in the ArcMap® 10.3 mapping software. The interpolated surface was sampled using a defined 2000 ft (610 m) spacing mesh and the extrapolated values were calculated using an inverse distance weighting function also available in the mapping software.

The fluid injection data were collected from publicly records made available by the TRRC which are annually reported as monthly injection volumes for all wells active in the state of Texas. For this model the injection volume data of all fluid injection wells injecting into the Ellenburger for which there are reported volumes as included. The TRRC public database contains the monthly injection volumes of these wells dating from 2005 up to the date of the last annual reporting date which is in September (Figure 3.3). To model injection activities, we discretized the 100 km^2 model space into 1 km^2 cells where the injection volumes of all wells located within each singular cell were combined in the model. The extent of the model space covering the northeastern portion of the Fort Worth Basin in the X and Y directions are X (3142000 to 3470000) and Y (-650000 to -322000) in the EPSG 32037 coordinate system which corresponds to the NAD27 Texas North state plane coordinate system. The extent of the model in latitude and longitude coordinates is 32.157°N to 33.057°N and 97.775°W to 96.707°W.

The fault geometries incorporated into the model were created by collaborators at TexNet and were created using a mixture of well log, seismic reflection, and earthquake location data (Hennings et al., 2019; Horne et al., 2020). The methodology for interpreting the fault geometries from these data sources is described in their publication. In this model we incorporate the fault interpretations defined as being “high-confidence faults” meaning that there is detailed

spatial and depth control for the fault geometries at the <5 km scale to interpret the extent, dip and segmentation of the faults. These high-confidence fault interpretations are in most cases associated with data collected from 3D seismic reflection profiles, densely spaced horizontal wells, and high-resolution earthquake catalogs based on local networks of stations. In total, 111 individual fault geometries were incorporated into our model associated with the northeastern portion of the FWB. Most of these fault geometries are basement rooted NE-SW trending normal faulting structures which in the eastern edge of the basin strongly parallel the Ouachita Thrust Front (Hennings et al., 2019; Horne et al., 2020). There is high variability in the distribution of fault structures present in the fault maps and incorporated into our model. These faults are associated with areas with variable availability of 3D seismic reflection and well log data for interpretation which may affect our modeling results. The details of the uncertainty associated with this issue are discussed in the later discussion section of this chapter (Figure 3.31). The fault geometries in the model were constructed from fault stick profiles for each fault, and each fault was given a 100 m wide damage zone due to the lack of data on the exact damage zone width of each individual fault. The stress tensor values along each fault were collected along each of the original fault stick segments at depth values corresponding to the top of the Ellenburger formation, the Ellenburger-Basement contact point, and ~2000 ft (~0.6 km) below the Ellenburger-Basement contact point.

3.9. Description of Model Realizations

Parameter testing of the model focused on three aspects: the eastern boundary condition of the model (which in prior FWB models is taken to be mapped at the intersection of the Ouachita Thrust Front and Ellenburger formation); the matrix permeability of the Ellenburger formation; and the spatial and magnitude distribution of faults incorporated into the model. A

suite of model realizations was prepared to explore how the range of different possible end-member cases affect the magnitude and the spatial distribution of the injection associated stress changes. The first set of end member realizations examines the effects of the eastern boundary condition of the basin on the model by running two versions of the model, one where the mapped contact point between the bottom of the Ellenburger formation and the Ouachita Thrust Front is treated as an impermeable feature and one realization where the boundary is treated as an open flow condition. The second set of model realizations examine the effects the bulk matrix permeability of the Ellenburger formation has on the model results using bulk permeabilities of 1, 10 and 100 mD. Lastly, a model realization was developed that included the direct connectivity of the fault networks along the eastern boundary of the basin. In all model realizations the faults are treated as highly permeable pathways allowing for the free flow of fluids within them. The complete list of model realizations including subset of results reporting peak pore fluid pressure increases across the entirety of the basin as well as peak pore fluid pressure increases at the site of Dallas-Irving sequence are recorded in Table 3.4. Based on preliminary knowledge of the behavior of pore fluid pressure and poroelastic stress diffusion it is expected that changes with regards to the conditions on the eastern boundary of the basin and the bulk matrix permeability value of the disposal formation should have the largest effects on the modeling results. Conversely, I expected that unless the total volume of each material layers defined as being a highly permeable fault structure is greatly increased the total change in the spatial and magnitude distribution of the injection associated stress changes will remain minimal.

3.10. Geomechanical Modeling Results

The injection associated stress changes documented in the modeling can be separated into two primary groups: the pore fluid pressure changes in the formations induced by the injection of

fluids; and the solid deformation due to poroelastic stress changes of those same affected formations which is represented as a stress tensor (Segall, 1989; Segall and Lu, 2015; Goebel et al., 2017; Keranen and Weingarten, 2018; Fan and Eichhubl, 2019). This combined stress change induced by injection activities in the basin can also be represented as a coulomb stress change value which is calculated using an assumed coefficient of friction, μ_f, and the calculated normal, σ, and shear stress, τ, components from the modeled stress tensor (Harris, 1998). The equation used to calculate this coulomb stress change is:

$$\Delta CFS = \tau + \mu_f(\sigma_n + p) \tag{3.7}$$

The basic characteristics of the modeling of fluid injection in the northeastern portion of the FWB which are not greatly affected by the different model realizations are described (Figure 3.32). First, both pore fluid pressure and poroelastic stress changes are concentrated in the portions of the basin in the immediate vicinity of the high-volume injection wells with peak values being located within Johnson County, the location of the Venus earthquake sequence site. Second, the poroelastic stress changes are lower in magnitude than the pore fluid pressure changes associated with the injection activities. Lastly, all injection associated stress changes are very low (<0.1 MPa) in the Dallas-Irving earthquake sequence area.

The results of the eastern boundary condition end-member model realization tests show that treating the boundary associated with the Ouachita Thrust Front feature as a nearly impermeable barrier leads to overall increases in the injection associated stress changes across the eastern margin of the basin (Figure 3.33). The question of how the eastern boundary condition of the basin should be treated stems from the question of how the Ouachita Thrust Front feature, or some associated system of fault features, interacts with the Ellenburger formation which is the main disposal formation in the basin. Prior models of the basin typically

take the mapped contact point between the Ouachita Thrust Front and the top or bottom of the Ellenburger (depending on the model) to represent a cessation in the eastward direction of the Ellenburger formation which precludes fluids from flowing east past this mapped boundary (Zhai and Shirzaei, 2018; Gao et al., 2021). However, there do exist some geologic cross-sectional interpretations across the mapped boundary taken in the Johnson County area that suggest some connectivity within the Ellenburger formation (Figure 3.8) (Magnani et al., 2017). Thus, we model two end-member conditions: first, a condition wherein there exists a continuous network of nearly vertical fault features, associated with the mapped Ouachita Thrust Front feature, which spans the entirety of the eastern boundary of the basin cutting through the Ellenburger formation and thus preventing fluid flow across the boundary; and second a condition where the Ellenburger remains continuous across the mapped boundary allowing for the free flow of fluids eastward. A comparison of these two model realizations is presented in Figure 3.32 where treating eastern boundary as an impermeable boundary forces the fluids and their associated stress changes to propagate further away from the injection wells in the north and south directions. The primary difference in the spatial distribution of the injection associated stress changes in these realizations is that when the fluids are allowed to flow freely throughout the model space, the pore fluid pressure and poroelastic stress changes follow a radial distribution pattern around each injection well. Meanwhile, the injection associated stress changes tend to concentrate and propagate along the eastern edge of the basin when that edge is treated as an impermeable feature. Thus, if we conclude that the model parameters should be defined in such a way that best match the interpreted geology of the basin and calculated average hydrologic and mechanical properties for each of the formations, then running the model with an impermeable eastern boundary like in prior models of the FWB is best (Figure 3.32).

The bulk matrix permeability of the Ellenburger has a strong effect on the overall magnitude and spatial distribution of the injection associated stress changes across the basin. Three model realizations were investigated with bulk matrix permeability values of 1, 10, and 100 mD for the Ellenburger formation (Figure 3.33). The need to define a bulk matrix permeability value for the Ellenburger formation units rather than using a complex heterogenous distribution of permeability values is due to limitations in the definition of parameters within the COMSOL software. When analyzing the permeability maps constructed for the FWB, and presented in Gao et al. (2021), it can be observed that the Ellenburger has pockets of higher and lower permeability values ranging from 1-1000 mD. It would require the definition and creation of thousands of individual formation materials and objects to match the complexity of these heterogeneous permeability maps within COMSOL. Thus, instead I chose to run separate model realizations using variable bulk matrix permeability values for the Ellenburger based on the distribution of values presented in the permeability maps to investigate how these different uniform permeability values affect the modeling results (Figure 3.33). Additionally, for each of the model realizations the permeability of the faults in the basin is taken to be 1000 mD making the faults more conductive high permeability pathways for which the fluids and their associated stress changes can traverse. The general relationship between the permeability of the Ellenburger formation and the magnitude and spatial distribution of the injection associated stress changes is that there is a tradeoff between the two. When examining the 1, 10, and 100 mD permeability cases we observe that the peak pore fluid pressure change observed within the model are around 8.34, 1.34, and 0.37 MPa for the respective models (Table 3.4). However, while the 1 mD Ellenburger permeability model does have the highest peak pore fluid pressure change observed in the basin, it comes with a tradeoff in the spatial distribution of the stress changes where the

fluids are restricted to the immediate regions surrounding the injection wells. This result is of special interest in order to examine the magnitude of the injection associated stress changes in the Dallas-Irving sequence area where the largest stress changes observed in that area occur in the 100 mD permeability model case. Examining the matrix permeability maps created by the BEG leads to conclusion that the model realization taking the bulk permeability to be 10 mD matches the well log data the best, and best matches the pore fluid flow model results created by the BEG (Gao et al., 2021). Temporal maps of each of the model realizations are shown in Figure 3.33 and the pore fluid pressure and poroelastic stress changes modeled at the Dallas-Irving sequence site are shown in Figure 3.34.

The final model realization tests focused on the effects of the fault network incorporated into the model on the magnitude and spatial distribution of the injection associated stress changes. The fault map incorporated into the model, which is based on a combination of well log, 2D and 3D seismic reflection data, and earthquake locations, represents the highest quality fault map available for the FWB (Figure 3.31) (Hennings et al., 2019). However, there are numerous parts of the basin for which there exists uncertainty about the number of faults in some areas due to a lack of seismic reflection data and earthquake locations in those regions. Two areas for which this is the case are the immediate region around the Dallas-Irving earthquake sequence and the eastern margin of the basin between the Dallas-Irving and Venus sequence sites. What interests me about the possibility of there being many missing faults is the stark contrast in the number of faults of various lengths and trends when comparing the fault map for northeastern Johnson County, for which 3D seismic reflection data was available, and the rest of the eastern edge of the basin (Hennings et al., 2019; Horne et al., 2020). Previous modeling tests showed that the bulk matrix permeability of each of the formations has a strong effect on the

overall magnitude and spatial distribution of the modeling results. In the models all faults are treated as high permeability features that cross between the various formations in the basin, therefore the number of faults incorporated into the model has a direct connection to the overall bulk matrix permeability of each of the formations (Figure 3.35). Therefore, if I assume that the fault distribution reflected by the 3D seismic reflection data is more accurate than the fault distributions created using solely by well log or 2D seismic reflection data, then even the updated fault maps of the basin underrepresent the total number of faults and thus underestimate the bulk matrix permeability of each of the formations in the models. To examine the effects of the input fault network map on the magnitude and spatial distribution of the injection associated stress changes, model realization tests were run where all faults were removed as an end member case designed to explore how the inclusion of the faults themselves affect the modeling results. These model realizations demonstrate that the inclusion of the network of faults has a noticeable effect on the spatial distribution of injection associated stress changes (Figure 3.35). The inclusion of the 111 faults greatly affects the bulk matrix permeability values for each of the formations. When running smaller model realization tests that only slightly altered the fault maps by including some small additional faults or increasing connectivity across the fault networks along the eastern edge of the basin, far smaller effects were found on the overall distribution of injection associated stress changes.

3.11. Discussion of the Geomechanical Modeling

Across the range of model realizations, the peak injection associated stress changes remain concentrated within the portions of the basin where the highest volume injection activities are taking place (Figures 3.32-3.35). While changing how the eastern boundary of the basin is treated or defining the bulk matrix permeabilities of the formations in the model and

including the fault maps as high permeability features do affect the modeling results; this general trend in the spatial distribution of the injection associated stress changes remains constant. Over the nearly 15-year period of documented injection the associated stress changes have propagated somewhat throughout the entire FWB. However, the overall magnitude of these far-field stress changes (>15 km from injection wells) remains far lower than the near-field stress changes. Regardless of what parameters are incorporated or altered in the model, only small (<0.1 MPa) injection associated stress changes are modelled at the site of the Dallas-Irving earthquake sequence (Figures 3.37). However, it should be noted that in each of those model realizations, except for the 1 mD Ellenburger formation permeability realization, that seismicity at the site of the Dallas-Irving earthquake sequence began only after the injection associated stress changes reached the fault. The model realization that best matches the calibrated bottom hole pressures from the Gao et al. (2021) model and the observed seismicity in the basin is the model realization, which incorporates a nearly impermeable boundary along the eastern margin of the basin, with the BEG provided fault network map and a bulk matrix permeability of 10 mD for the basin (Figure 3.33).

The modeling results described below show that injection activities are occurring within the Ellenburger and that faults exist throughout the basin that hydraulically connect the Ellenburger and crystalline basement formations; that pore fluid pressure changes lead to instability in both the Ellenburger and crystalline basement while poroelastic stress changes stabilize the Ellenburger and de-stabilize the crystalline basement. The pore fluid pressure changes within the model are highest within the Ellenburger formation and as noted previously are concentrated near injection well locations. Additionally, while pore fluid pressure changes related to the injection activities are observed throughout the basin, the overall magnitude of the

stress changes quickly decrease at distances >15 km from the injection well sites (Figure 3.37). The pore fluid pressure changes are highest within the Ellenburger formation and propagate into the crystalline basement using the permeable fault features as pathways. The increase in pore fluid pressure de-stabilizes the faults and encourages slip in both the Ellenburger and crystalline basement formations. However, the poroelastic stress changes associated with the fluid injection activities behave differently as they stabilize and strengthen the Ellenburger formation while also de-stabilizing the underlying crystalline basement formation (Figure 3.38). The overall magnitude of the poroelastic stress changes is lower than that of the pore fluid pressure changes (on the order of 1/10th) within the model. However, this pattern of poroelastic stress changes stabilizing the Ellenburger may be a factor as to why seismicity in the FWB is concentrated within the basement formation rather than the Ellenburger formation.

The models created here can be directly compared to other models examining injection associated stress changes in the FWB, such as those created by Zhai and Shirzaei (2018) and Gao et al. (2021). The primary differences between these three different studies of the FWB relate to the input data sets incorporated into the models and what specific injection associated stress changes each model sought to calculate. The Zhai and Shirzaei (2018) model, like the one presented here, is a coupled geomechanical model that calculates both pore fluid pressure and poroelastic stress changes. However, this earlier model was created prior to the major advancements of our understanding of the geology and geometry of the basin and so lacks information about the dipping nature of the major formations in the basin and the updated fault maps with a far greater number of faults present throughout the basin (Hennings et al., 2019; Smye et al., 2019; Horne et al., 2020). Meanwhile, the Gao et al. (2021) model incorporates largely the same input data sets as the model created here; however, their model is a pure fluid

flow model which only calculates the pore fluid pressure changes and does not incorporate poroelastic stress changes. What can be observed when comparing the results of each of the models is that the end time results of each models show very similar spatial distributions in the injection associated pore fluid pressure changes (Figure 3.35). However, what is different across the three models is the overall magnitude of the injection associated stress changes across the basin. The Zhai and Shirzaei model places peak pore fluid pressure changes at approximately 4 MPa, while the Gao model records peak basin-wide stress changes of only ~2 Mpa, and the model presented here predicts a peak pore fluid pressure change of ~1.5 Mpa. These differences are most likely due to the input parameters used in each model. As stated earlier the Shai and Shirzaei model used simplified geometries of the basin formations due to a lack of refined data about each of the formations. The differences in the new COMSOL model and Gao model are likely due to two key factors: one the new model incorporates poroelastic stress changes which alter how the full tensor of the injection associated stress changes is calculated; and second, that while I do use ostensibly the same input parameters, the COMSOL© software we used to run these models differ in how those input parameters are incorporated. The model realization tests reveal that the bulk matrix permeabilities of each of the formations most strongly controls the overall magnitude and spatial distributions of the injection associated stress changes (Figure 3.33). The Gao model was created using the Petrel© software which allowed for the definition of highly heterogeneous distributions of formation parameters such as matrix permeabilities. However, the new models were created using the COMSOL© modeling software which only allows the definition of a bulk matrix permeability value for each material. This limitation means that to mimic the heterogeneous nature of the matrix permeabilities presented in the Gao model, one would need to define 10s of thousands of different materials with their own complex

geometries and defined geologic and hydrologic parameters (Gao et al., 2021). Such an approach quickly became unfeasible to accomplish within the COMSOL modeling software and thus a bulk matrix permeability value for the entirety of each of the formations using the distributions presented in the Gao et al. (2021) publication was used and most likely led to discrepancies in our end modeling results. Overall, all three models describe injection associated stress changes that remain most highly concentrated in those regions of high-volume injection activities which over longer time periods diffused outwards across the basin (Figure 3.35).

3.12. Discussion of the Dallas-Irving Sequence

The poroelastic modeling results wherein the injection of fluids leads to a greater stability within the Ellenburger disposal formation, but instability within the underlying crystalline basement formation matches the expected results explored in prior modeling studies (Segall and Lu, 2015; Chang and Segall, 2016). Many different induced seismicity sequences across the world have been associated with poroelastic stress change effects over the years (e.g., Segall et al., 1994; Deng et al., 2016; Goebel et al., 2017; Zhai et al., 2019; Chang and Yoon, 2020). These case studies and theoretical modeling studies focus on poroelastic stress changes as the driving factor of induced seismicity due to poroelasticity operating on a much faster time scale and being farther reaching than direct pore fluid pressure diffusion (Goebel et al., 2017; Goebel and Brodsky, 2018). The theoretical models created by Chang and Segall (2016) show that when a fault hydraulically connects a fluid reservoir and the underlying basement formation, both pore fluid pressure and poroelastic stress changes migrate from the reservoir into the basement with the pore fluid pressure changes being higher magnitude than the poroelastic stress changes. In both my FWB model and their theoretical model the same behavior of the poroelastic stress changes is observed, wherein the reservoir formation where injection is occurring becomes more

stable while the underlying basement formation becomes unstable and thus more prone to slipping (Figure 3.38; Chang and Segall, 2016). Spatially, poroelastic stress changes are important as they are farther reaching and capable of inducing seismicity on faults at greater (>15 km) distances from injection wells (Goebel et al., 2017), such as the Dallas-Irving sequence. My model does show that poroelastic stress changes can be observed throughout the basin even at locations in the basin with no nearby injection wells. However, the magnitude of the poroelastic stress changes in my model are very small (~0.01 MPa) and thus unlikely to be able to induce seismicity on their own. Additionally, the pore fluid pressure changes at the Dallas-Irving site are also higher than the poroelastic stress changes, which may be due to some measure of direct pore fluid pressure diffusion from the injection wells closest to the sequence, the two wells at the DFW airport (Figure 3.1a), reaching the fault. Based on the modeling results, the injection associated stress changes reached the Dallas-Irving fault in early 2008 and have only steadily increased in magnitude since then (Figure 3.34).

The template matching results suggest that seismicity associated with the Dallas-Irving fault began sometime in late 2013 (Hennings et al., 2021) or early 2014 (this study) based on which station data set is used. Local monitoring of the Dallas-Irving sequence began only in January 2015, but the USGS catalog does contain several events throughout 2013-2014 that may have occurred on the Dallas-Irving fault but due to higher location uncertainties were not located on the now defined fault plane. Nonetheless, the modeling results show that injection associated stress changes reached the Dallas-Irving fault prior to the onset of seismicity regardless of which estimated start date of the sequence is used. In all three model cases the injection associated stress changes recorded at the Dallas-Irving sequence site remain small in magnitude, however in each model this injection associated stress changes are modeled to have reached the fault

location prior to the onset of seismicity. This means that while the injection associated stress changes at the Dallas-Irving site are small (<0.1 MPa) they still precede the fault becoming seismically active and thus we conclude that the Dallas-Irving sequence was induced by the cumulative oil and gas-related activities, particularly waste fluid disposal, ongoing within the FWB.

3.13. Conclusions

The Dallas-Irving earthquake sequence is unique among the FWB sequences due to the lack of nearby injection wells within a major metropolitan area and the generation of a M3.5 and M3.6 doublet many months after the sequence began. The sequence occurs along a NE-SW trending normal fault with most seismicity being located within the pre-Cambrian crystalline basement formation. While the sequence lacks a nearby injection well that might provide a direct pore fluid pressure increase, the sequence still has many of the key characteristics of an induced seismicity sequence. A coupled geomechanical model of the FWB created to examine the stress changes from basin-wide injection revealed that while these injection activities had associated stress changes that were small at the Dallas-Irving sequence site, the stress changes did reach the fault prior to the onset of seismicity. These models also revealed that while modeling input parameters affect the magnitude and spatial distributions of these stress changes, the overall characteristic spatial distribution of the injection associated stress changes remains consistent with the largest stress changes observed in those areas of high-volume injection activities. *EQcorrscan* template matching codes were applied to 6.5 years of seismic data to better assess when seismicity at the Dallas-Irving sequence began. The *EQcorrscan* template matching codes are cross-correlation based event detectors which utilized the NTXES Dallas-Irving earthquake catalog to produce templates to help find missing events in the dataset. The template matching

codes revealed that seismicity associated with the Dallas-Irving fault may have begun as early as February 2014 using the waveform data collected by stations active in the FWB. The resulting template matched event detection list also contains information about specific phase arrival times associated with each event detection which when extracted could be utilized with separate event association and location algorithms to transform the detection list into a full earthquake catalog. Ultimately, I conclude based on the modeling and template matching results that the Dallas-Irving sequence is an example of a fluid injection induced earthquake sequence.

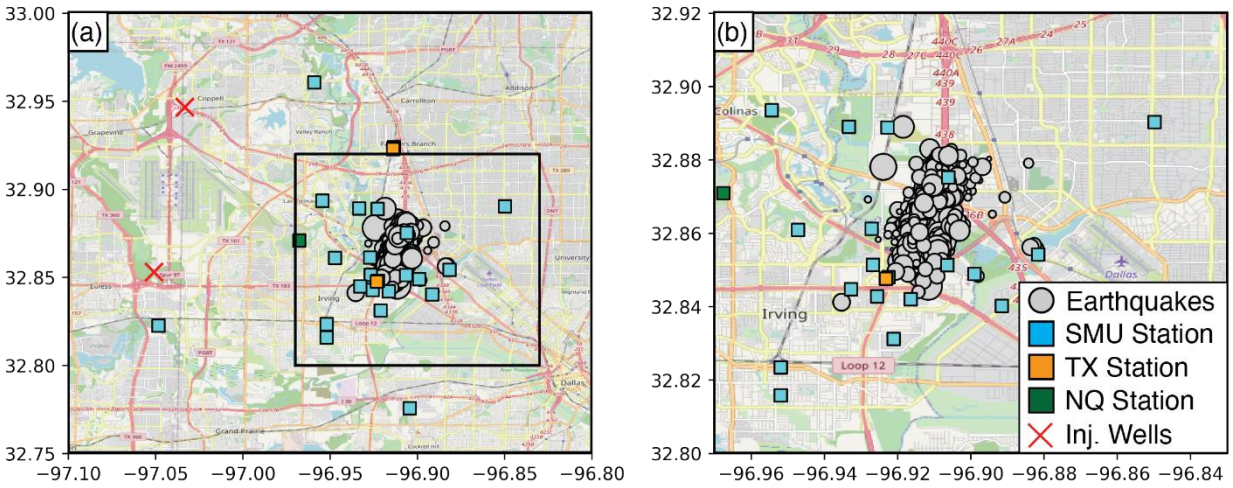

Figure 3.1. (a) Map view of the Dallas-Irving earthquake sequence scaled by magnitude with local and regional monitoring stations shown along with the locations of the nearest fluid injection wells. The black box represents the map view area used in panel (b). (b) Zoomed in view of the Dallas-Irving earthquake sequence focusing on the only the local monitoring stations around the sequence.

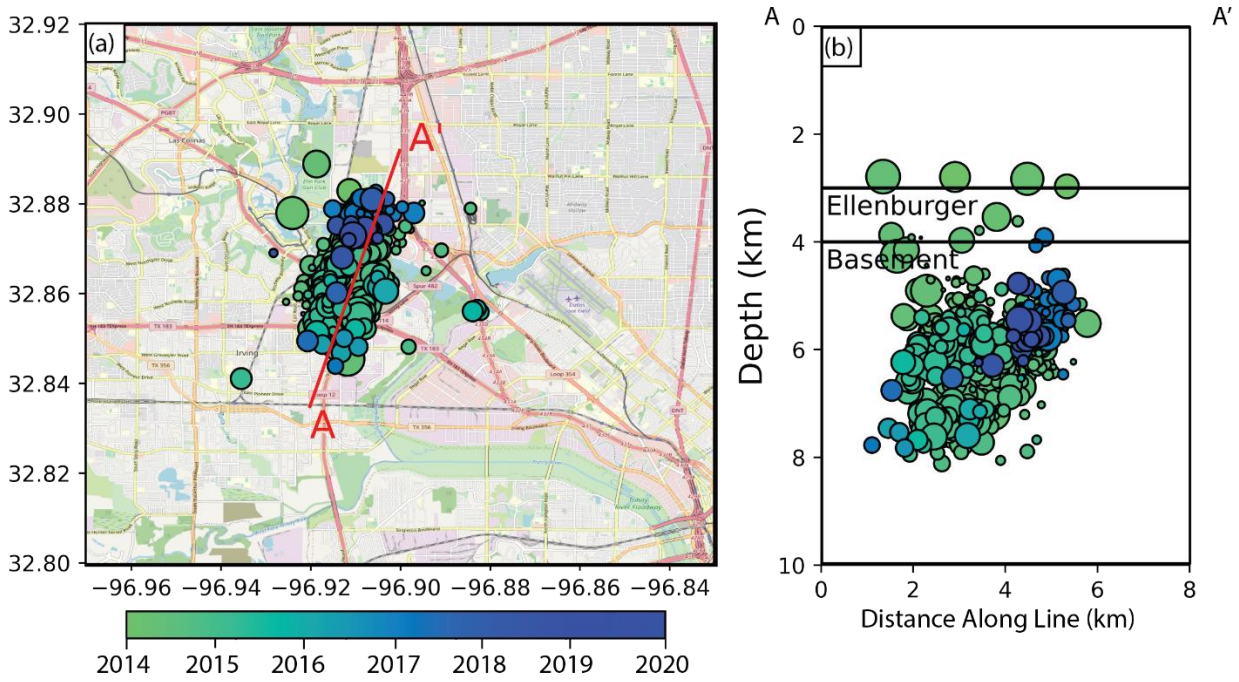

Figure 3.2. (a) Map view of the Dallas-Irving earthquake sequence with size scaled by magnitude and colored by each event's origin time. Also shown is the line denoting the length that the depth cross-sectional view covers (red line) along with the area for which events were plotted within the cross section (dashed lines). (b) Cross-sectional view of the Dallas-Irving earthquake sequence with the same size-magnitude and color-origin time scaling as used in the map view (a). Also shown are the formation top depths of the limestone-dolomite Ellenburger and crystalline Basement units for this portion of the Fort Worth Basin.

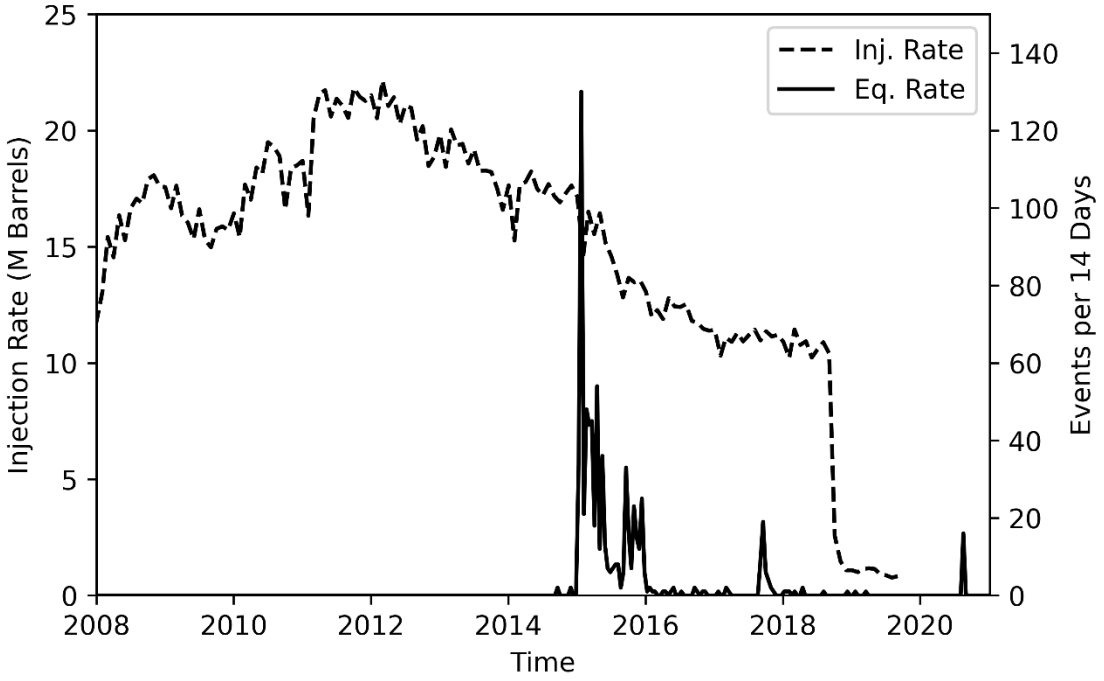

Figure 3.3. Time history of fluid injection and seismic activities occurring within the entirety of the Fort Worth Basin and the Dallas-Irving earthquake sequence, respectively. The fluid injection values are reported as monthly injection volumes in millions of US Barrels (dashed line). The Dallas-Irving earthquake rate is reported as the number of events occurring every 14 days (solid line).

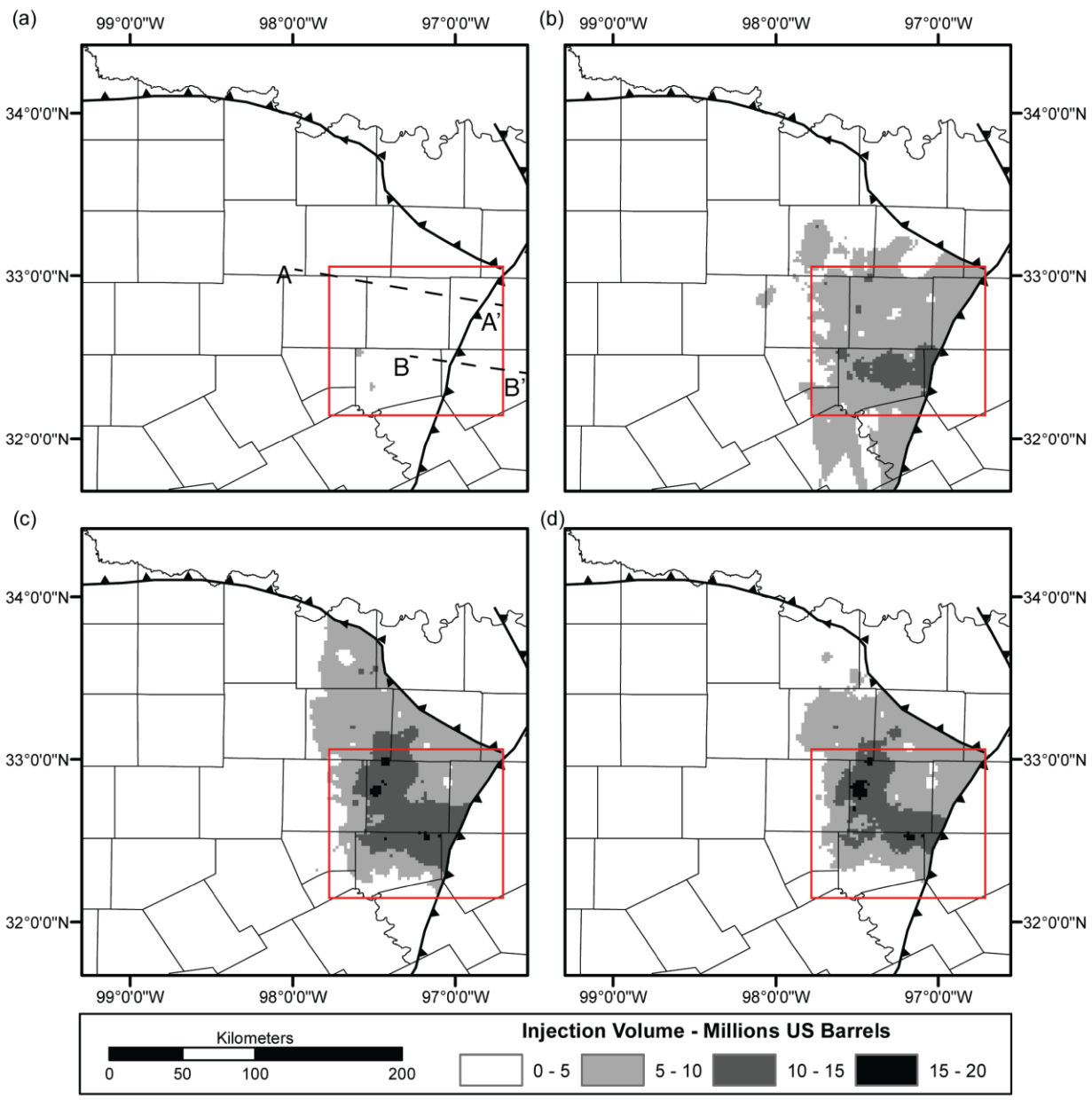

Figure 3.4. Map views of the volumes of injected fluids in the FWB over the time periods of (a) 2005–2007, (b) 2008–2010, (c) 2011–2013, and (d) 2014–2017 with the boundary of area covered by the coupled geomechanical model shown by the red box. Also shown in panel (a) are the cross-sectional lines noting the geologic cross sections used for figures 3.7 and 3.8.

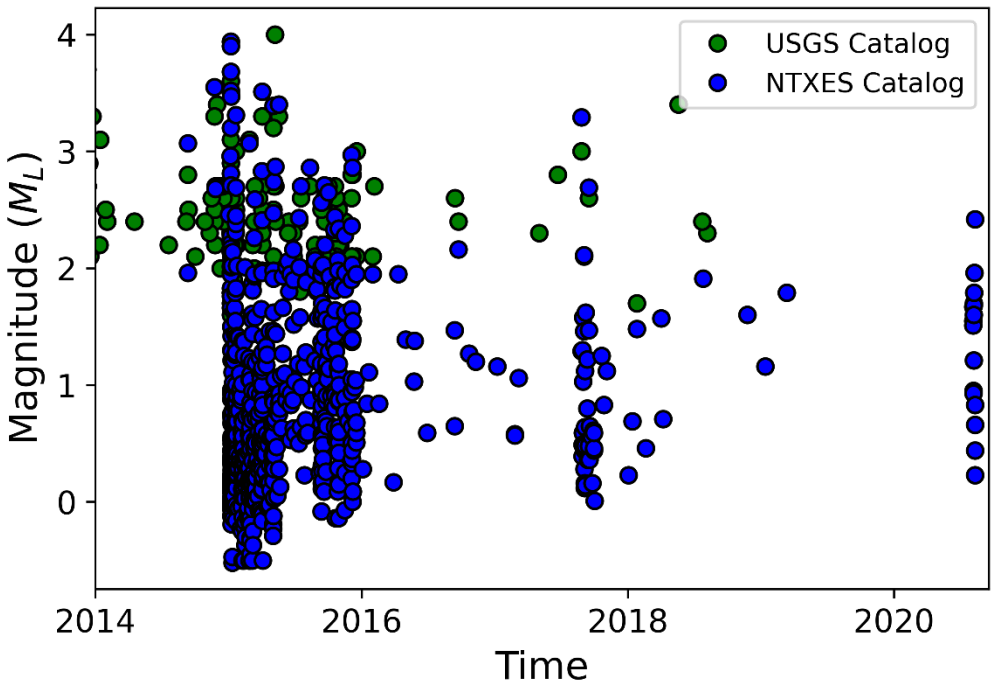

Figure 3.5. Magnitude vs. time plot of the Dallas-Irving earthquakes present in the NTXES catalog (blue circles) and the events recorded within the USGS earthquake catalog (green circles) in the Fort Worth Basin over that same period.

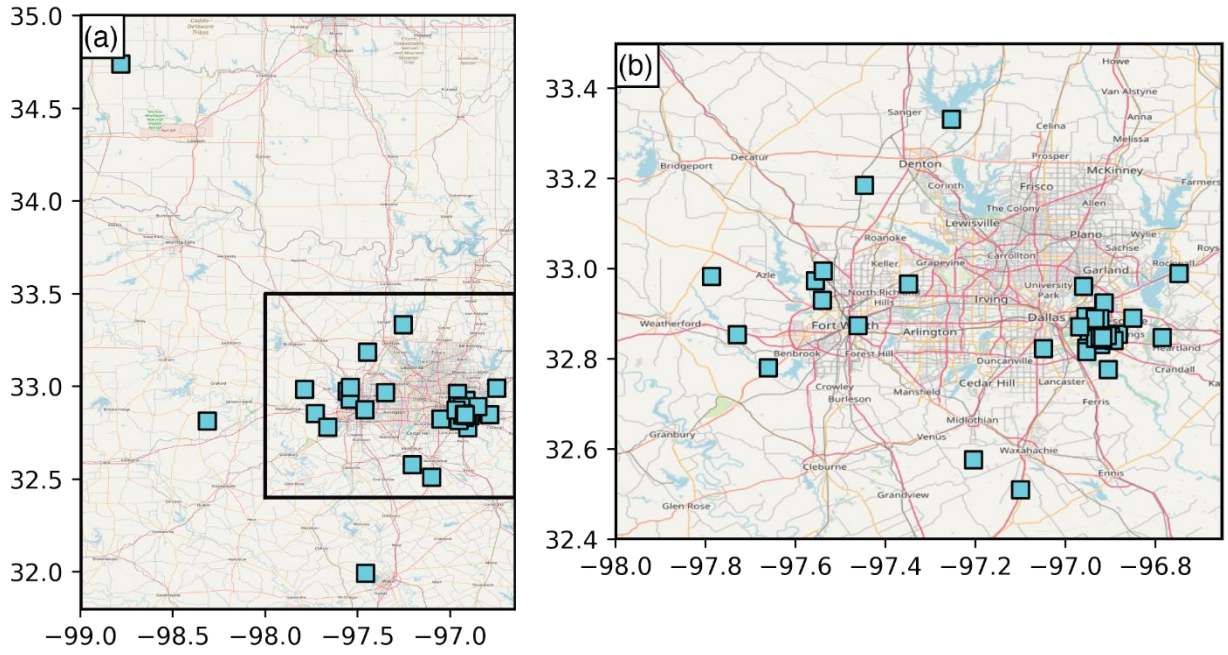

Figure 3.6. (a) Map view of the stations (cyan squares) used within the template matching process with the black box denoting the map area used for panel (b). (b) Zoomed in view of the network of stations used in the template matching process for the Dallas-Irving earthquake sequence.

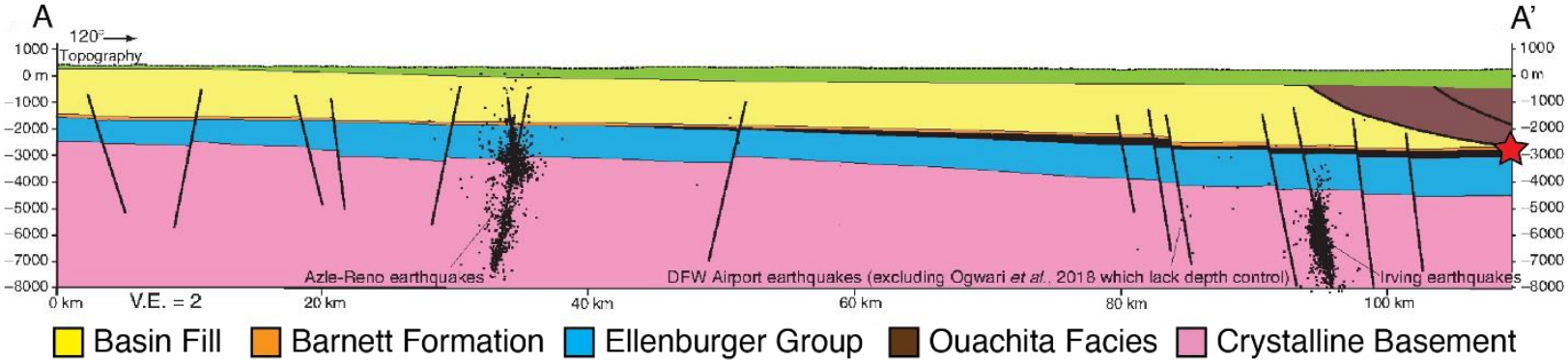

Figure 3.7. Geologic cross-section taken from the northern portion of the Fort Worth Basin covering the cross-sectional A-A' line shown in Figure 3.4. The red star denotes the mapped contact point between the Ouachita Facies and the Ellenburger formation which for modeling purposes is considered the eastern edge of the hydraulically connected portion of the Ellenburger formation. This geologic interpretation was originally presented in the publication Hennings et al., (2019).

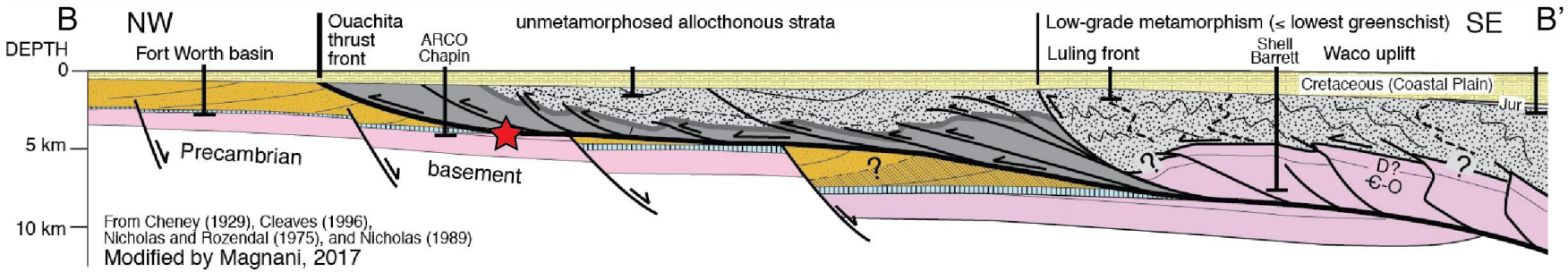

Figure 3.8. Geologic cross-section taken from the southern portion of the Fort Worth Basin covering the cross-sectional B-B' line shown in Figure 3.5. The red star denotes the mapped contact point between the Ouachita Facies and the Ellenburger formation which for modeling purposes is considered the eastern edge of the hydraulically connected portion of the Ellenburger formation. This geologic interpretation was originally presented in the publication Magnani et al., (2017).

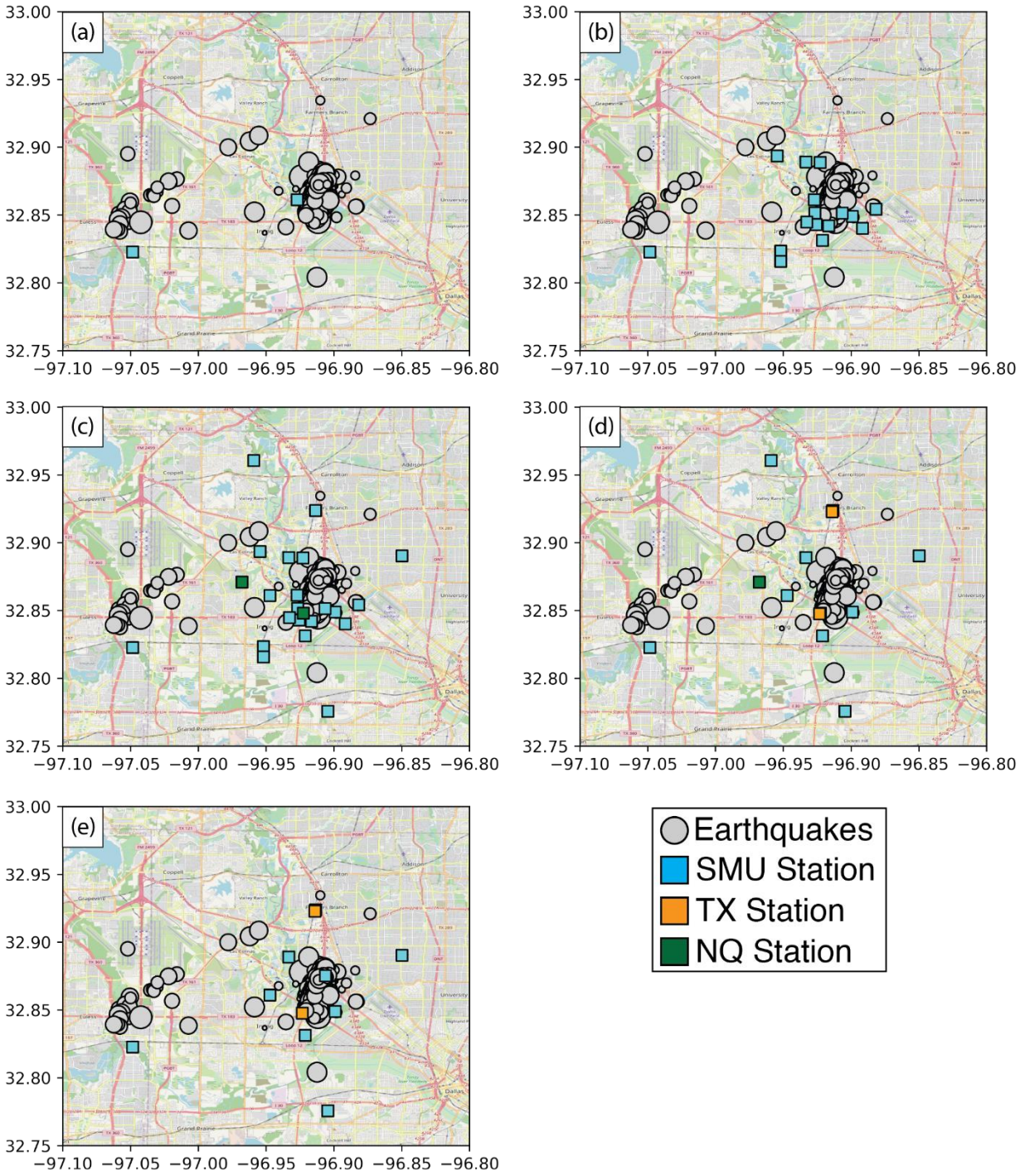

Figure 3.9. Map views of the NTXES earthquake catalog with the stations used in the template matching process shown as colored squares following the color scheme shown in the legend for the time periods of (a) 2014-2015, (b) Jan. 1st – Jan 7th 2015, (c) Jan. 7th – Dec. 31st 2015, (d) 2016-2017, and (e) 2017-2021.

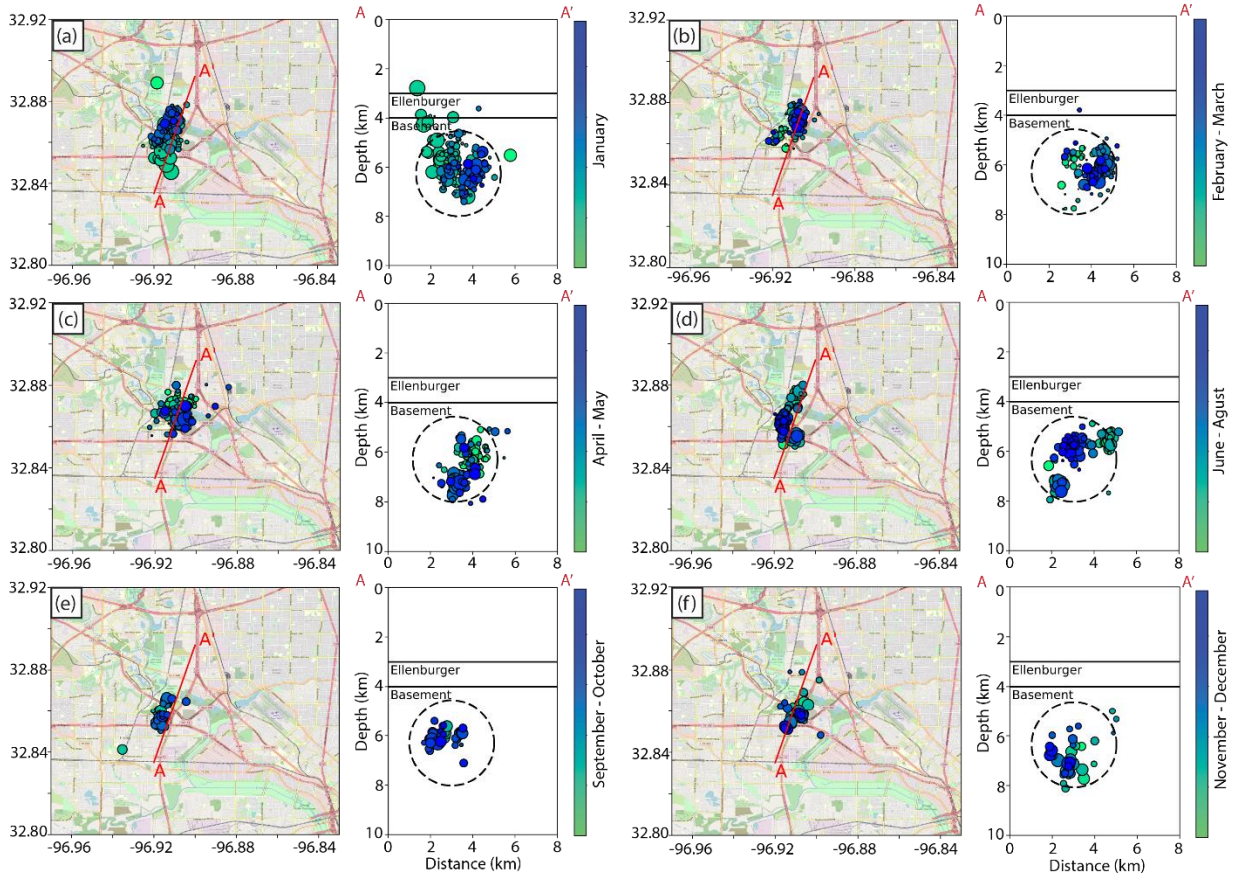

Figure 3.10. Map and cross-sectional depth views covering the 2015 period NTXES catalog earthquakes scaled by magnitude separated into views of (a) January, (b) February – March, (c) April – May, (d) June- August, (e) September – October, and (f) November – December. The cross-sectional views cover the mapped A – A' line in each map view (red line). The tops of the Ellenburger and crystalline basement formations are shown in each depth cross section along with a circular patch denoting the active fault zone.

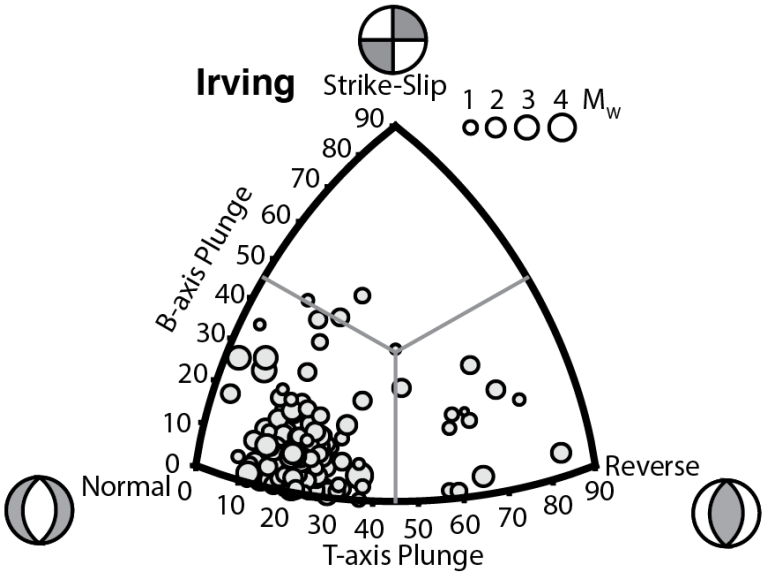

Figure 3.11. Ternary plot of the Dallas-Irving focal mechanisms (gray circles) scaled by the magnitude of each event.

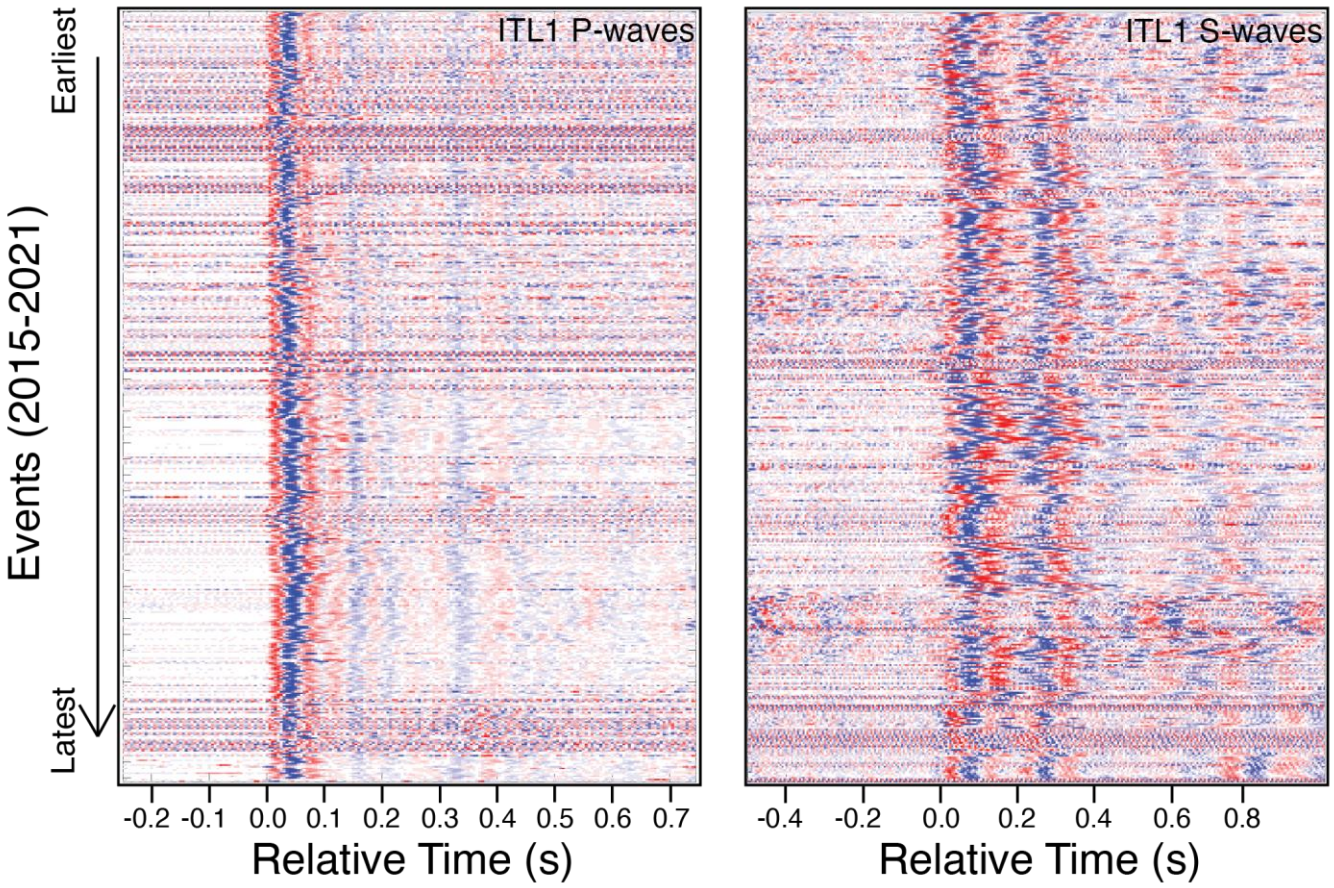

Figure 3.12. Plots showing the shaded versions of all the waveforms of detected P-waves (left) and S-waves (right) on station ITL1 from 2015-2021 with positive values shown in blue and negative values in red. All waveforms are plotted relative to their recorded phase arrival time which is denoted as 0.0 s.

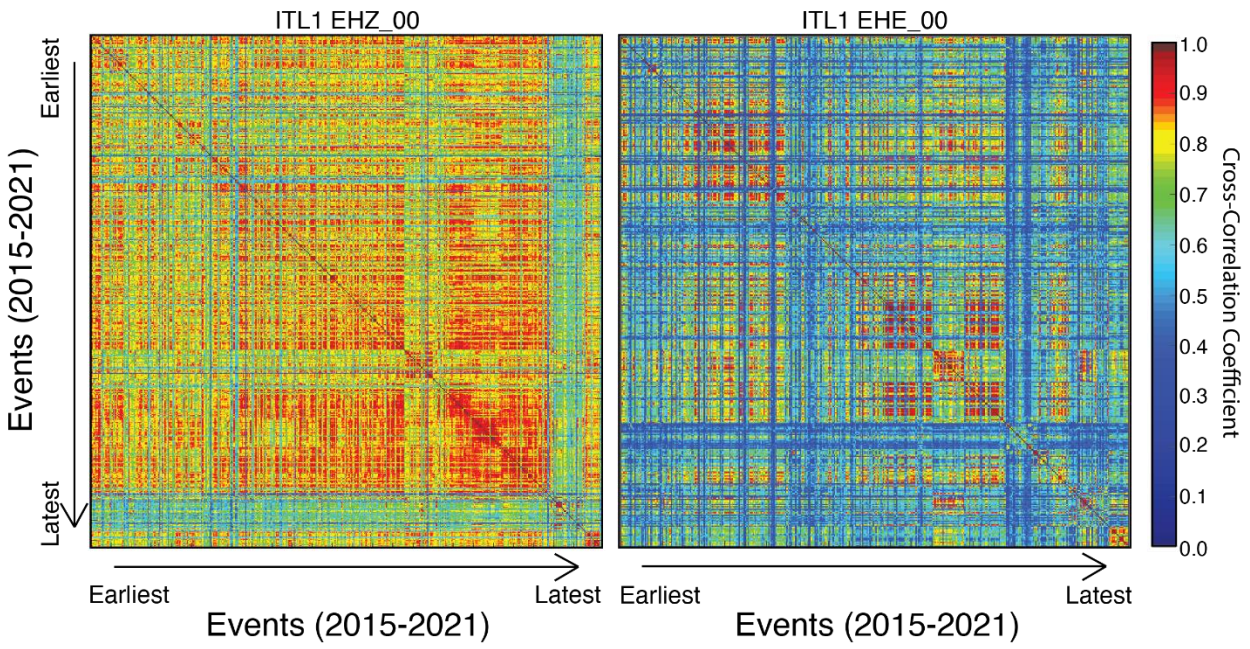

Figure 3.13. Matrices showing the event-to-event cross-correlation values for all events recorded on station ITL1's vertical (left) and east horizontal (right) channels. The diagonal line of events with cross-correlation values of 1.0 represents when an events was cross-correlated with itself.

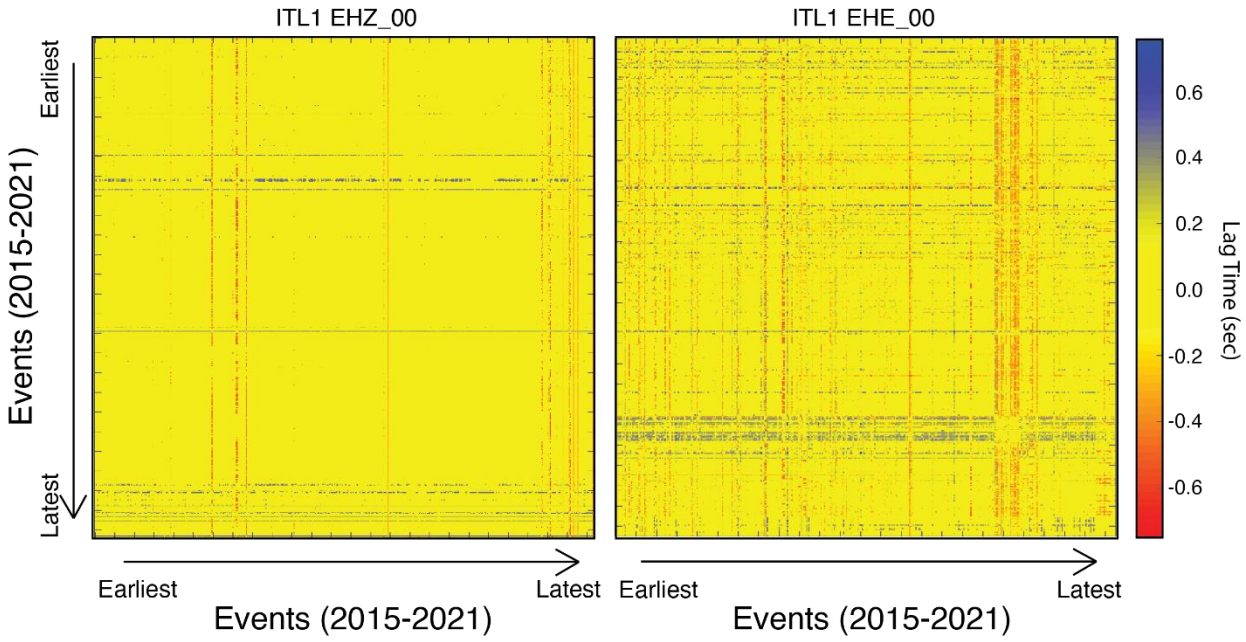

Figure 3.14. Matrices showing the event-to-event lag time values calculated from the cross-correlation results for all events recorded on station ITL1's vertical (left) and east horizontal (right) channels. The lag times are reported in seconds.

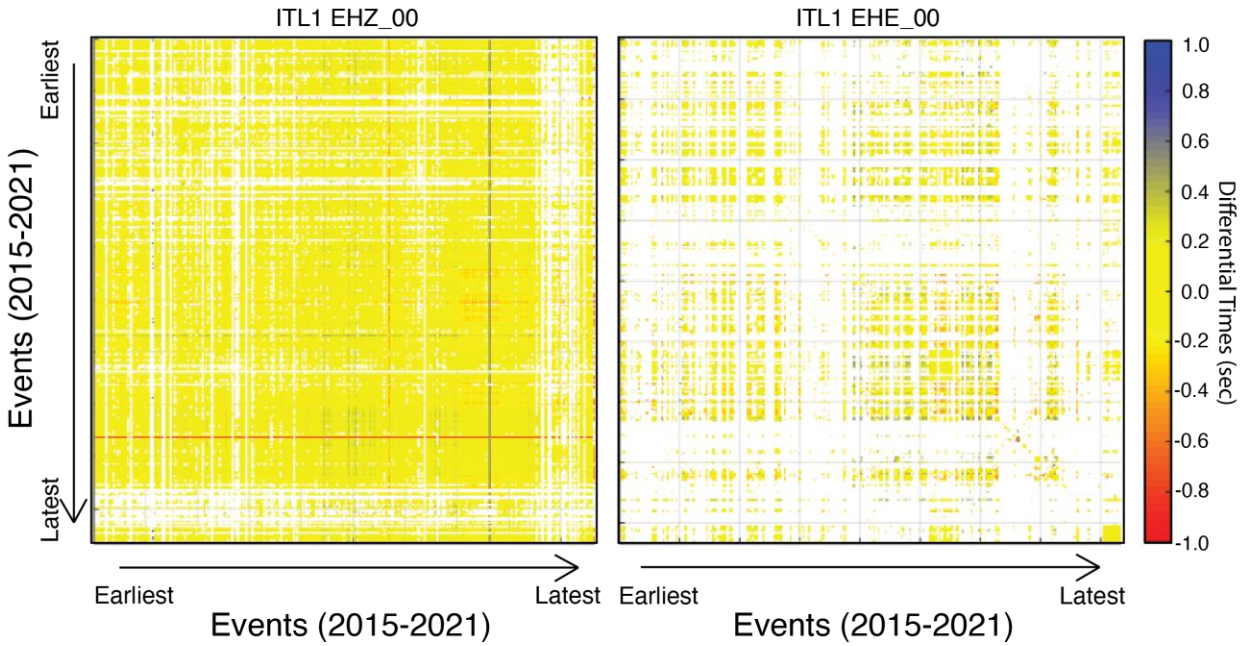

Figure 3.15. Matrices showing the event-to-event differential time values for all events recorded on station ITL1's vertical (left) and east horizontal (right) channels. Only the differential time values for event pairs with a minimum cross-correlation value of 0.6 are shown in the plots. The differential time values are reported in seconds.

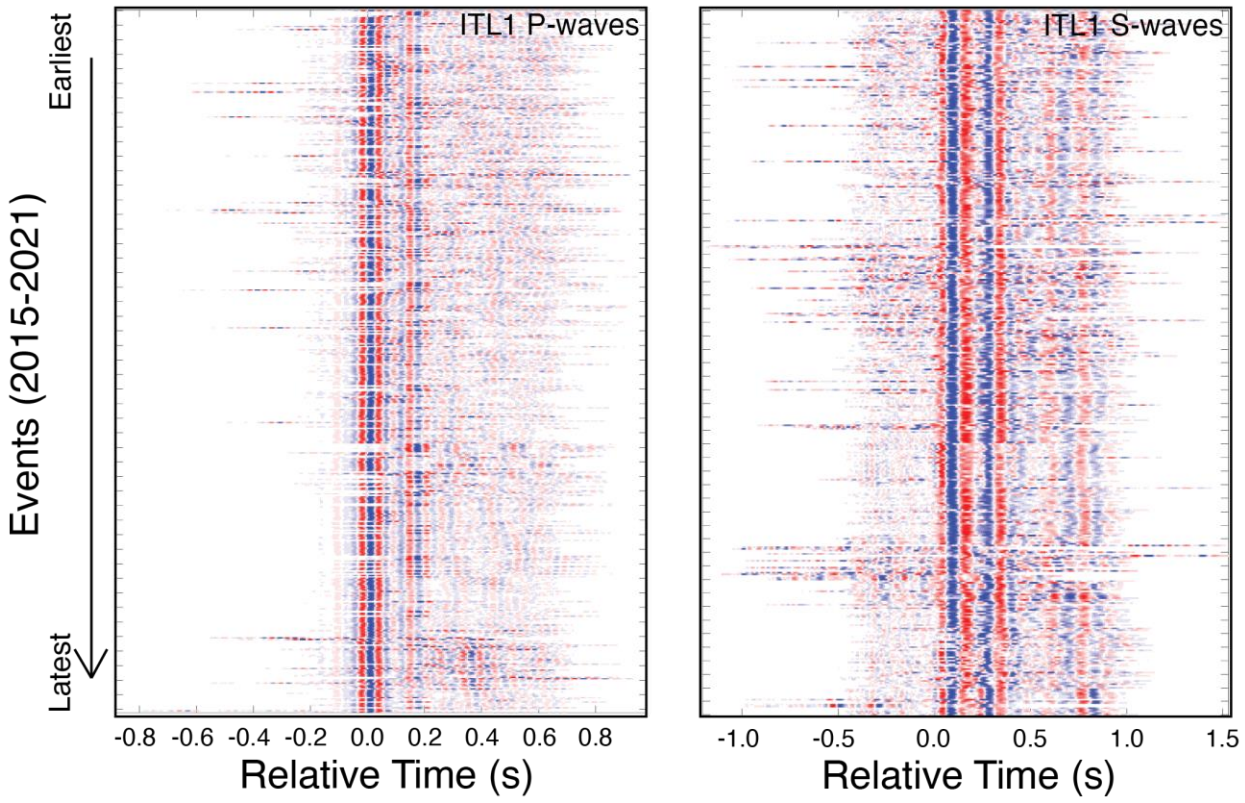

Figure 3.16. Plots showing the shaded versions of all the lag time corrected waveforms of detected P-waves (left) and S-waves (right) on station ITL1 from 2015-2021 with positive values shown in blue and negative values in red. The relative time value of 0.0 s is taken to be lag time corrected phase arrival time for each event.

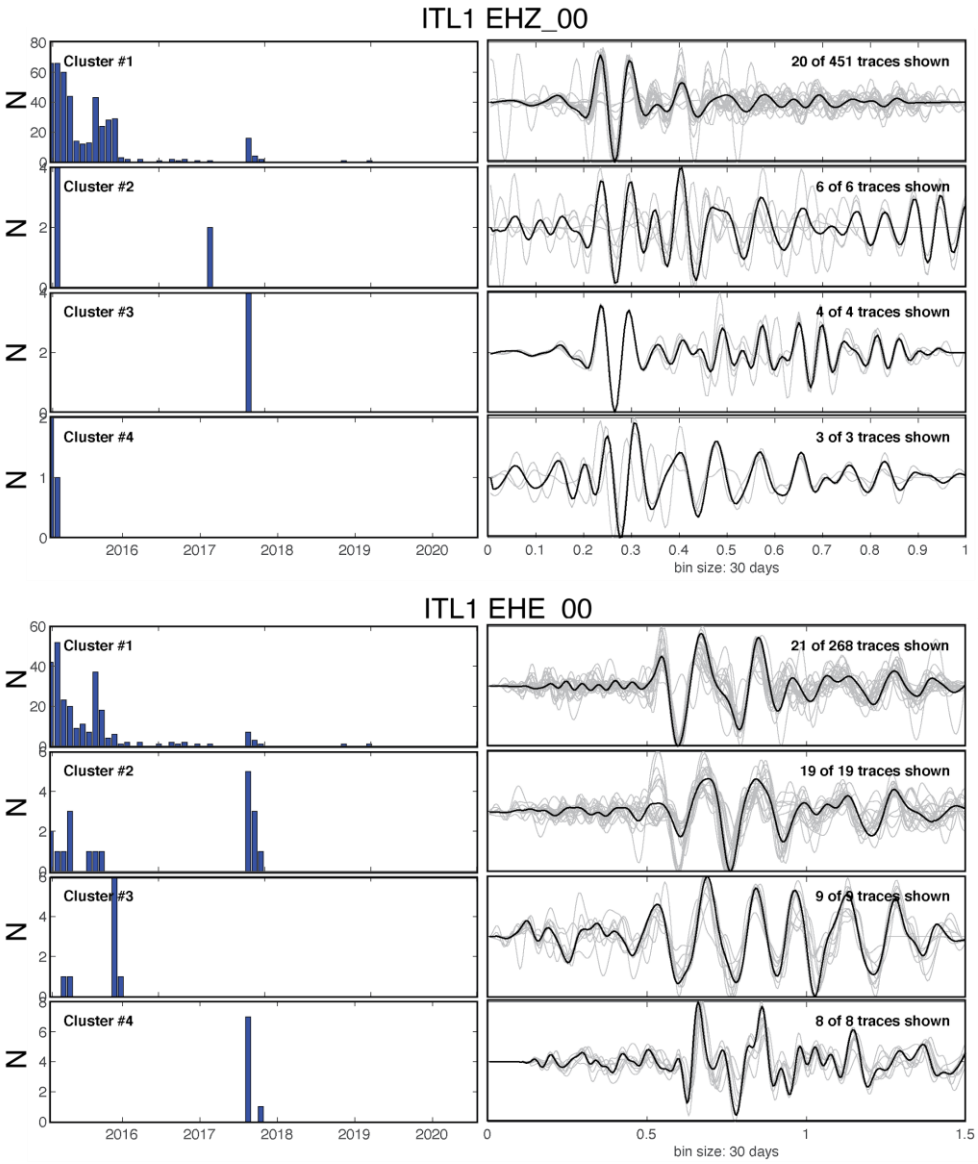

Figure 3.17. Plots showing identified clusters of events with similar waveforms detected on the vertical (top) and east horizontal (bottom) channel based on a minimum cross-correlation value of 0.6. The left column shows a histogram view of events associated with each cluster over time using a bin size of 30 days. The right panel shows the average waveform signal for each cluster in black and a sample of the waveforms associated with each cluster in light gray. The number of associated waveforms shown for each cluster is noted in the waveform panel.

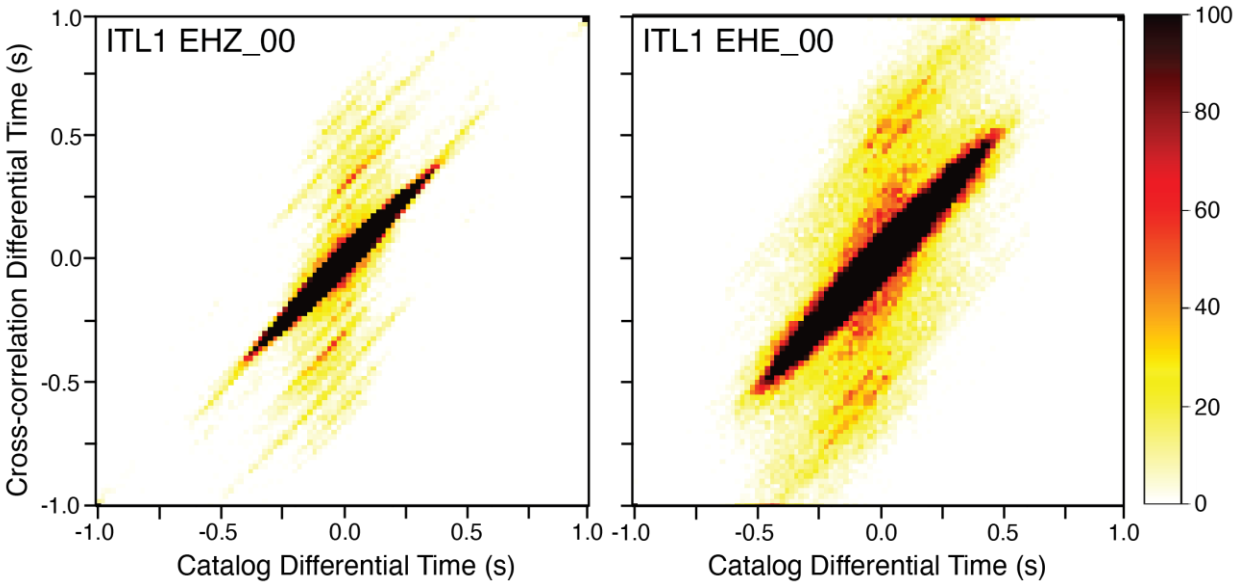

Figure 3.18. Heat maps of the catalog and cross-correlation differential time values for those events detected on ITL1's vertical (left) and east horizontal (right) channels.

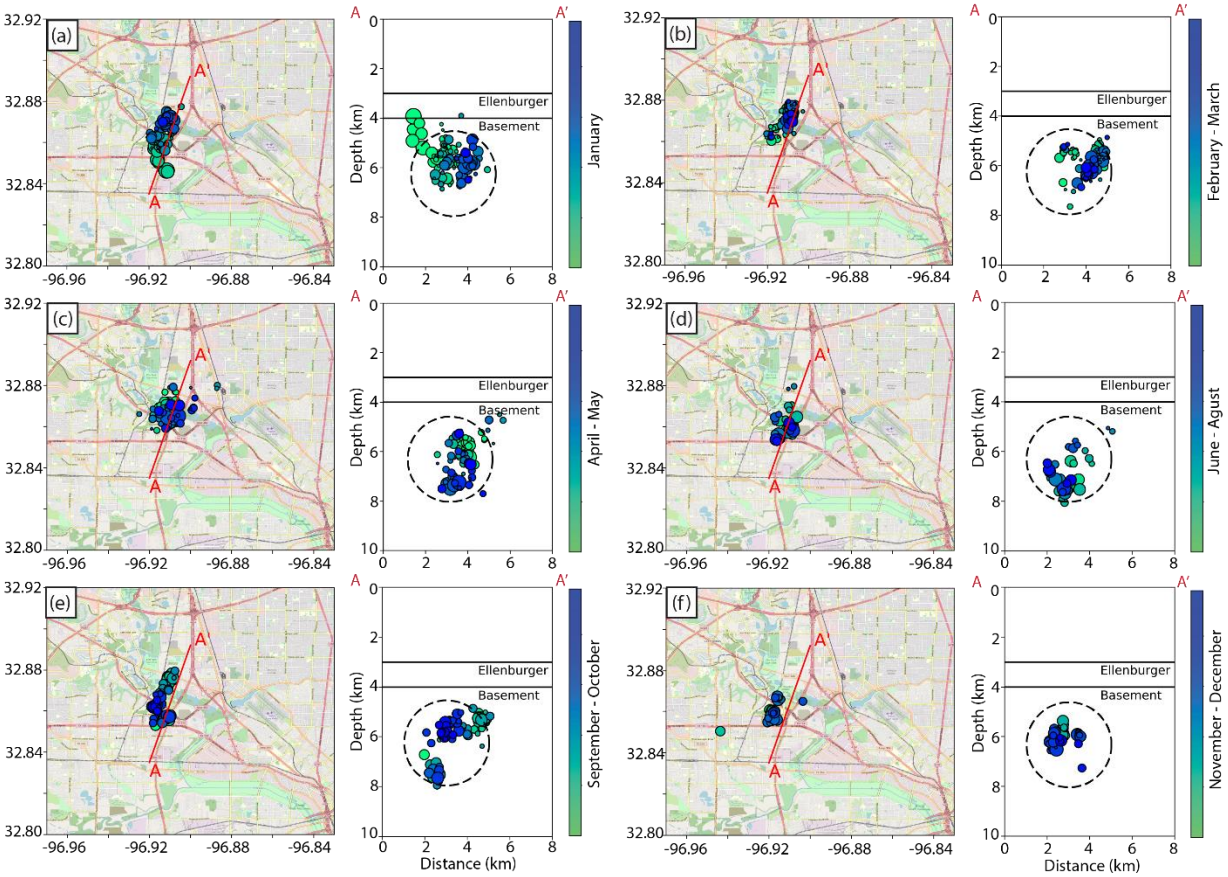

Figure 3.19. Map and cross-sectional depth views covering the 2015 period double-difference catalog earthquakes scaled by magnitude separated into views of (a) January, (b) February – March, (c) April – May, (d) June- August, (e) September – October, and (f) November – December. The cross-sectional views cover the mapped A – A' line in each map view (red line). The tops of the Ellenburger and crystalline basement formations are shown in each depth cross section along with a circular patch denoting the active fault zone.

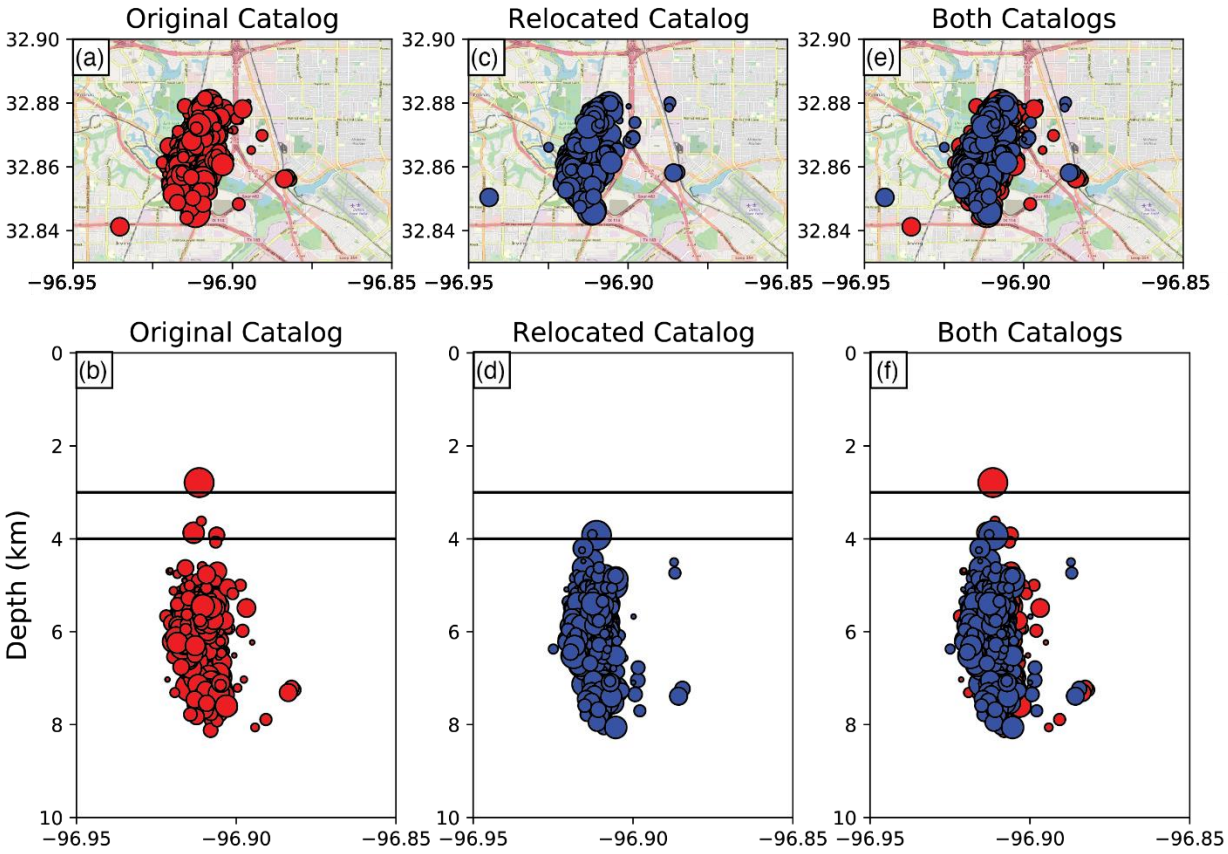

Figure 3.20. Map and depth cross-sectional views with the earthquake symbol sizes scaled by magnitude for the original NTXES Dallas-Irving earthquake catalog (a,b), the double-difference relocated Dallas-Irving earthquake catalog (c,d), and both catalogs plotted together (e,f). The lines in the depth cross-sectional views (b,d,f) represent the top of the Ellenburger (3 km) and crystalline basement (4 km) formations.

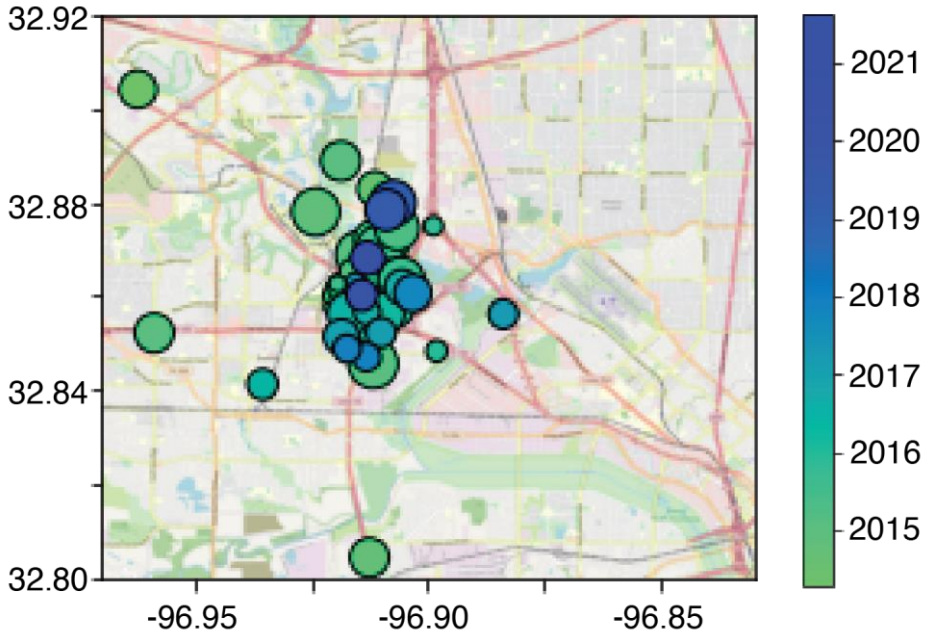

Figure 3.21. Map view of the 51 template events used within the template matching process for the Dallas-Irving earthquake sequence with their size scaled by magnitude and their color by the origin time of the event.

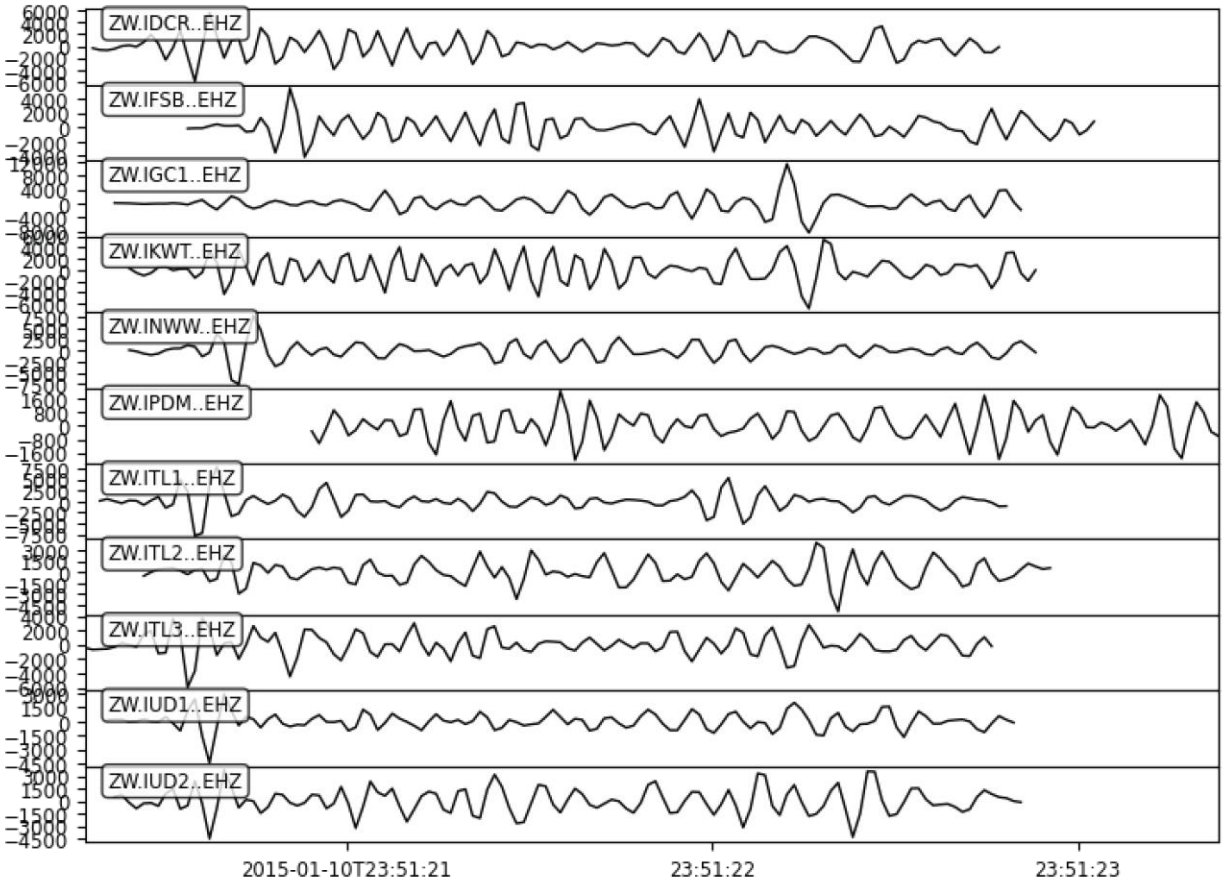

Figure 3.22. Waveforms collected from a template event used within the template matching process which occurred on Jan. 10th, 2015 at approximately 23:51:20 UTC time and was recorded using the displayed stations. The waveforms for this template have been processed using a bandpass filter of 10-40 Hz and have had their trends and means removed.

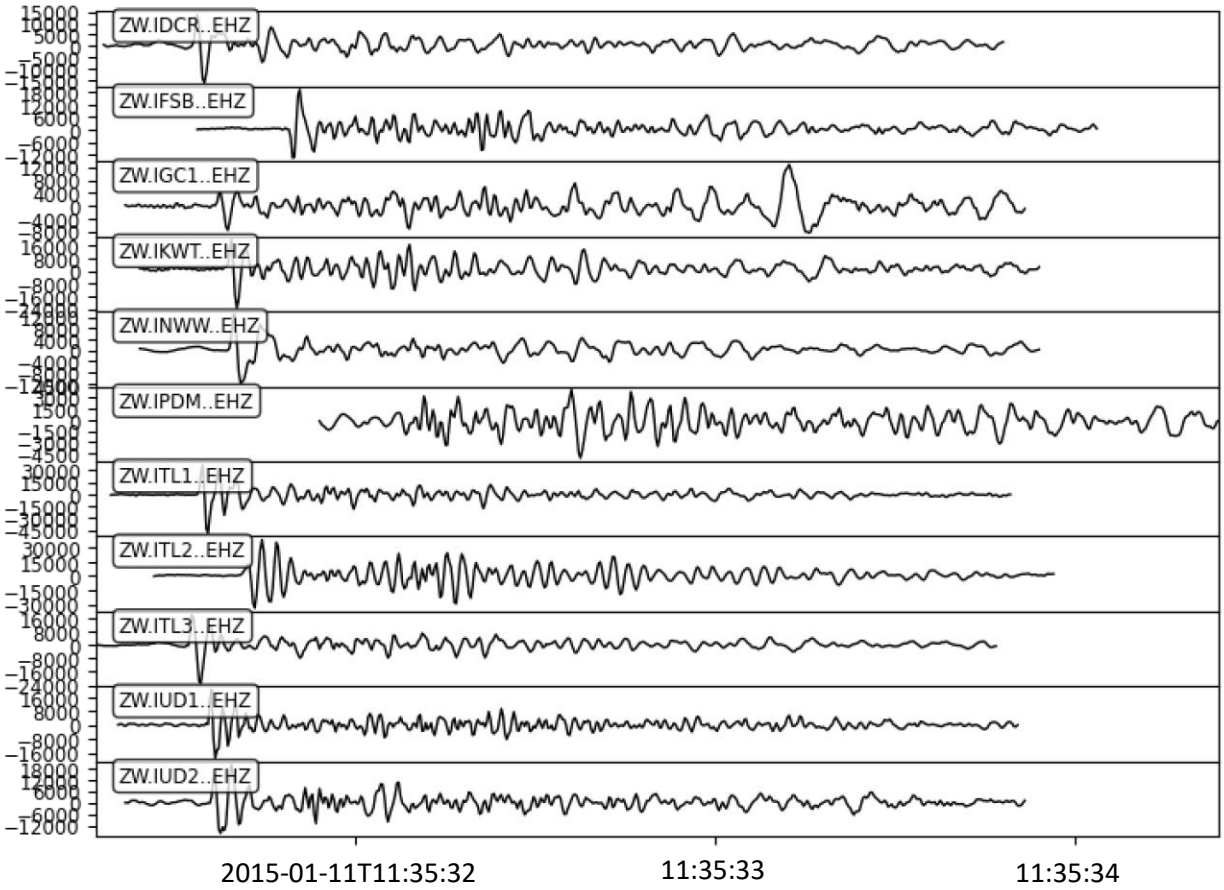

Figure 3.23. Waveforms for an event detected using the template shown in Figure 3.22 which was detected on Jan. 11th, 2015 at approximately 11:35:31 UTC time. The waveforms shown have not been processed.

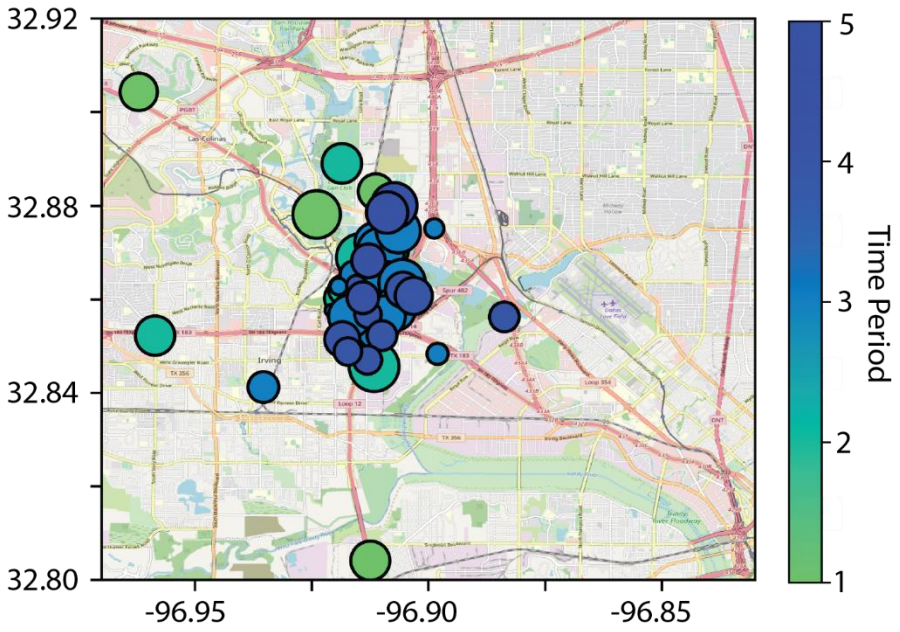

Figure 3.24. Map view of the 51 template events used in the template matching process with the symbol sizes representing the event magnitudes and the colors representing which of the five designated time periods each of the templates occurred in. The five time periods are as follows: (1) pre-2015, (2) Jan. 1st – Jan. 7th 2015, (3) Jan 7th – Dec. 31st 2015, (4) 2016-2017, and (5) 2017-2021.

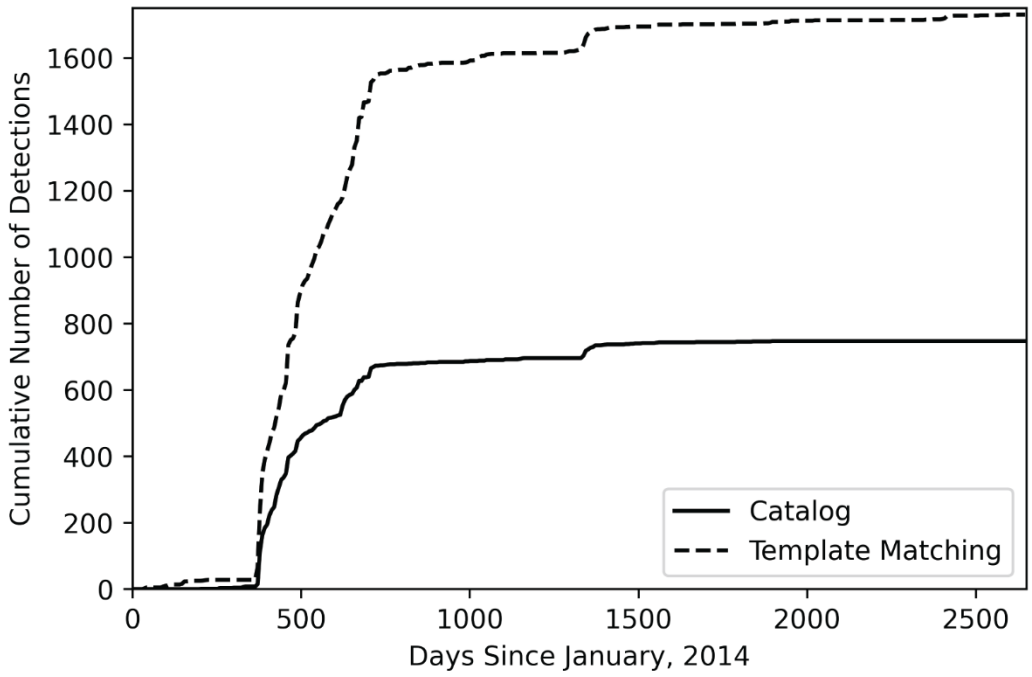

Figure 3.25. Temporal view of the number of events detected over time within the NTXES earthquake catalog associated with the Dallas-Irving fault (solid line) and using the *EQcorrscan* template matching approach (dashed line).

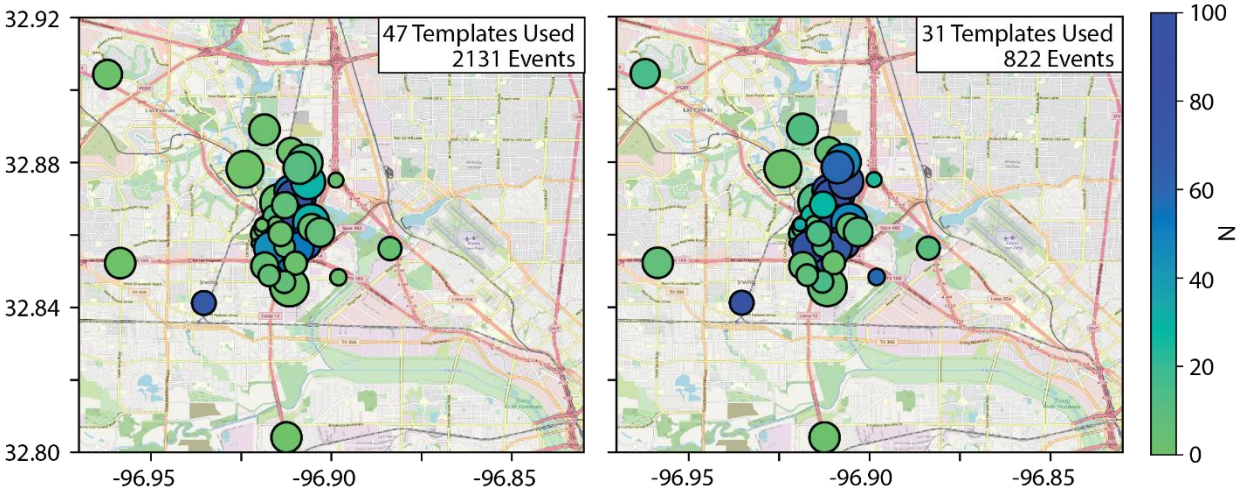

Figure 3.26. Map views of the templates which had associated event detections using the *EQcorrscan* template matching approach with the symbol sizes representing the magnitude of the template event and the color representing the number of events detected using that template. The two maps show the full list of events detected which had at least one cross-correlated phase pick assigned (left) and the filtered list of event detections which had at least eight cross-correlated phase picks assigned (right).

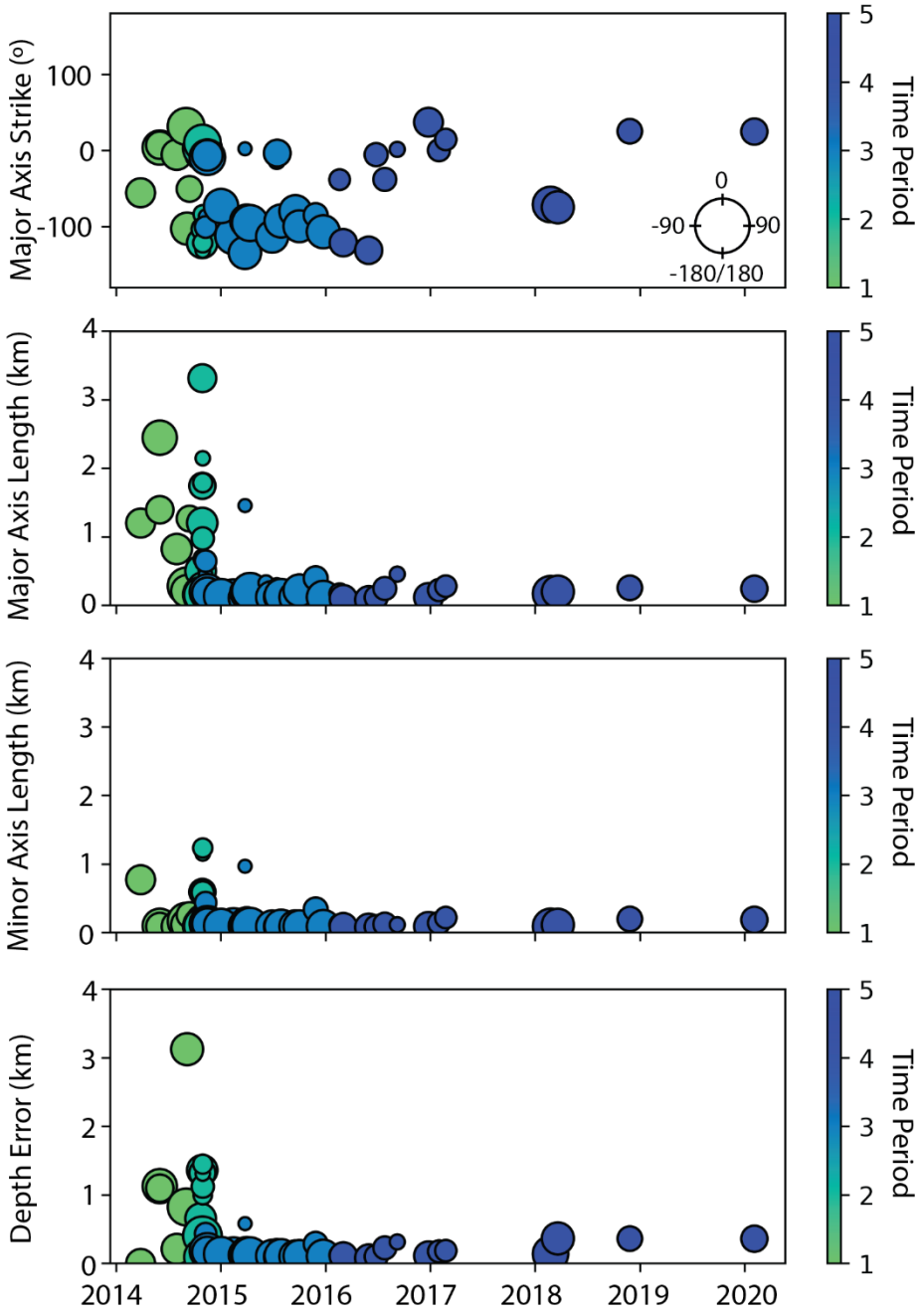

Figure 3.27. Temporal plots showing the location uncertainties associated with each of the template events with the symbol sizes representing the template event magnitude and the color representing which defined period each template is associated with. The location uncertainties are reported as 68% confidence interval error ellipsis with the strike of the major axis being shown in degrees relative to north with west defined as -90° and east as 90°. The five time periods are as follows: (1) pre-2015, (2) Jan. 1st – Jan. 7th 2015, (3) Jan 7th – Dec. 31st 2015, (4) 2016-2017, and (5) 2017-2021.

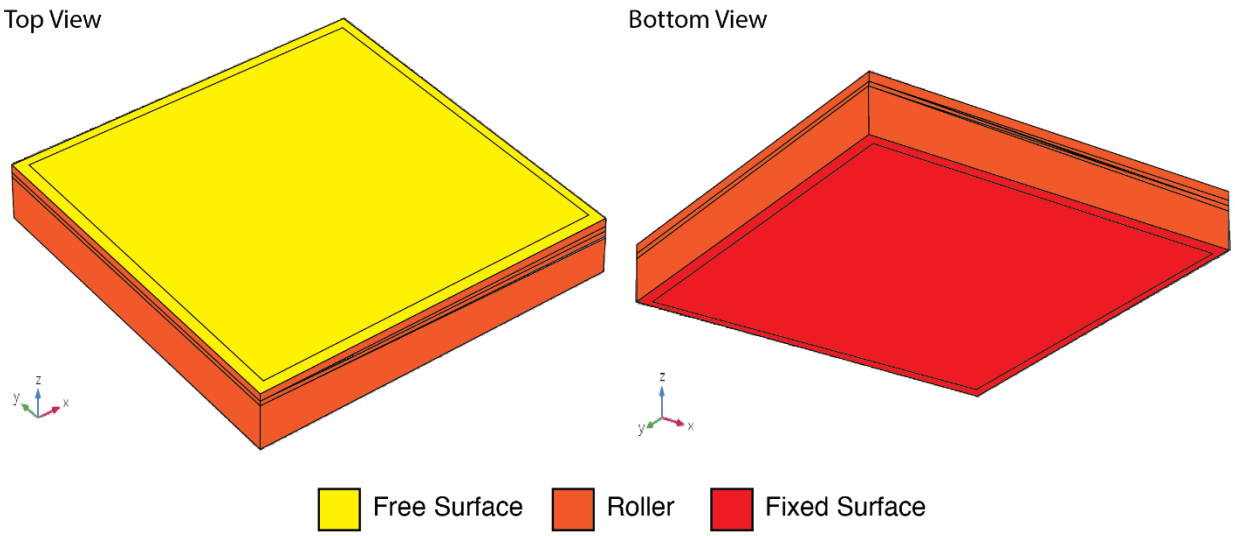

Figure 3.28. 3D views of the COMSOL geomechanical model showing a top (left) and bottom (right) view of the model. The outer boundaries of the model are colored to denote the boundary conditions imposed by the Solid Mechanics module upon the model. The axis orientations for each view of the model are also shown with the y-axis representing the north direction.

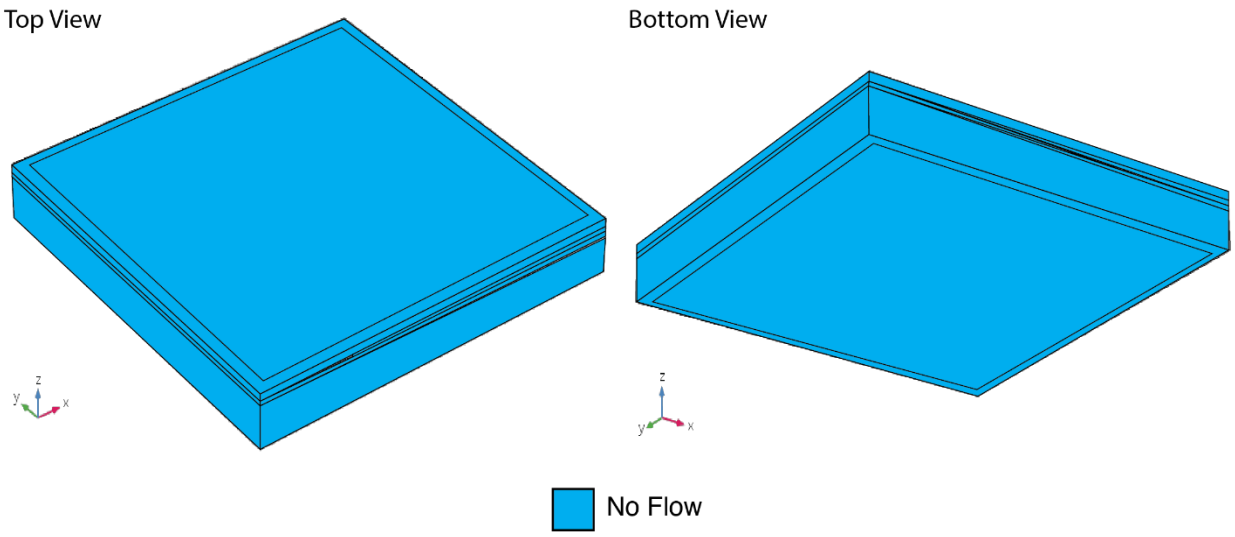

Figure 3.29. 3D views of the COMSOL geomechanical model showing a top (left) and bottom (right) view of the model. The outer boundaries of the model are colored to denote the boundary conditions imposed by the Darcy's Law fluid flow module upon the model. The axis orientations for each view of the model are also shown with the y-axis representing the north direction.

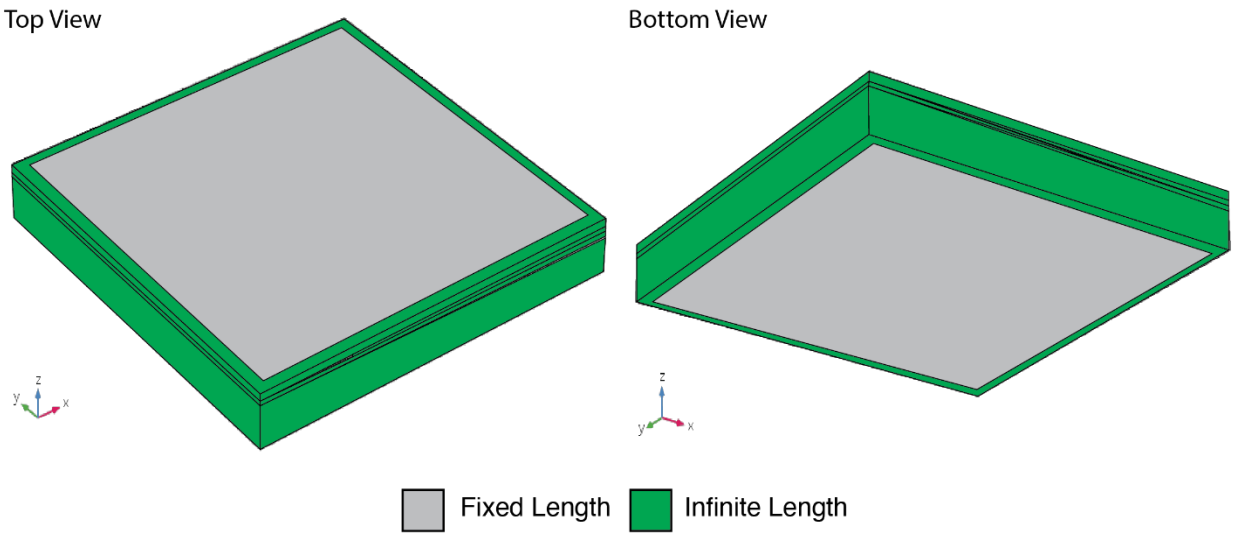

Figure 3.30. 3D views of the COMSOL geomechanical model showing a top (left) and bottom (right) view of the model. The outer bounding area of the model which is treated as having an infinite length when compared to the defined area of the inner portion of the model is highlighted in green. The axis orientations for each view of the model are also shown with the y-axis representing the north direction.

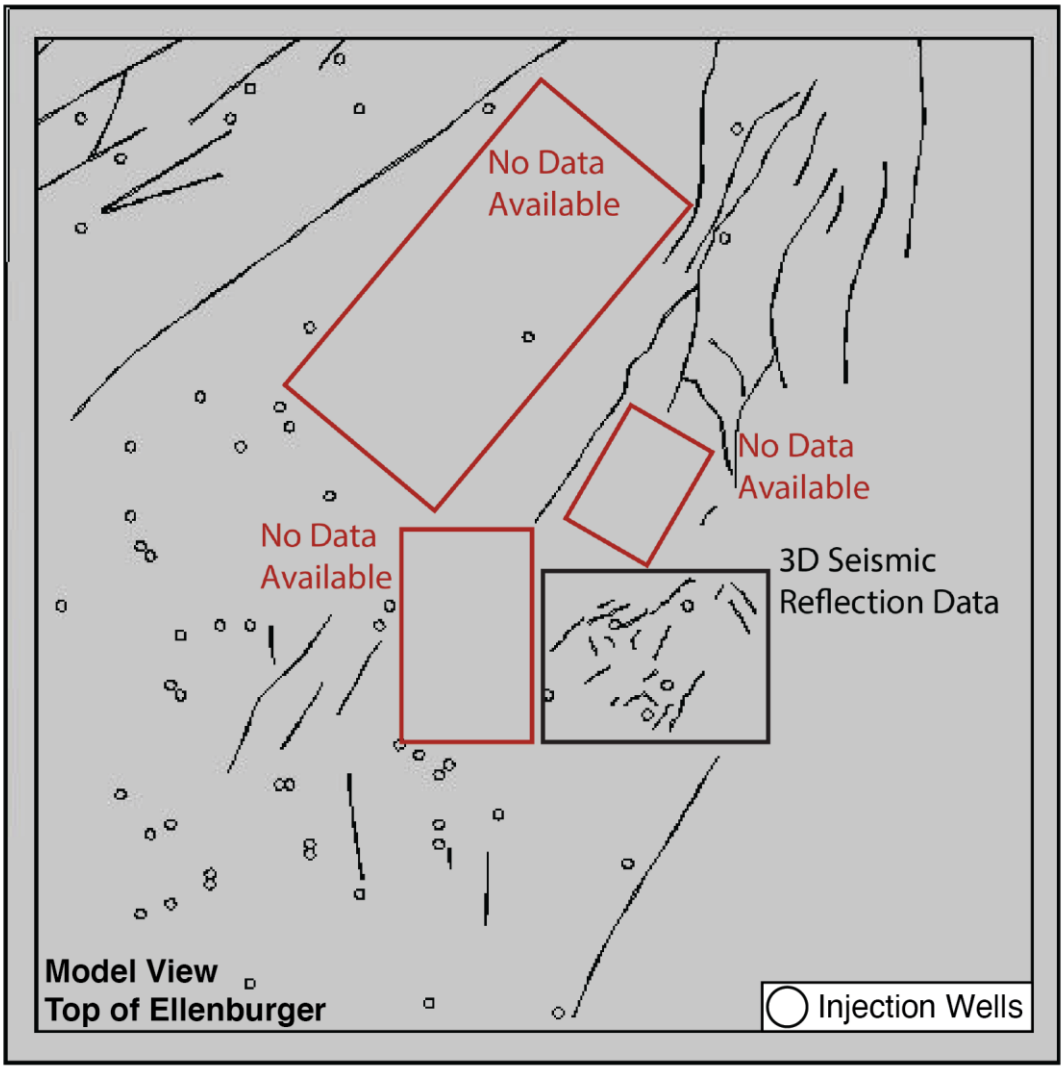

Figure 3.31. View of the COMSOL geomechanical model from the top of the Ellenburger formation showing the locations of the injection wells in the model as circles and the faults as linear features. Also denoted are the areas within the model space for which there was no available data that could be used to attempt to interpret fault geometries from (labeled red boxes) along with the portion of the model space where 3D seismic reflection data was available (labeled black box). The other portions of the basin where faults are mapped were interpreted using a mixture of available well log, 2D seismic reflection, and earthquake catalog data.

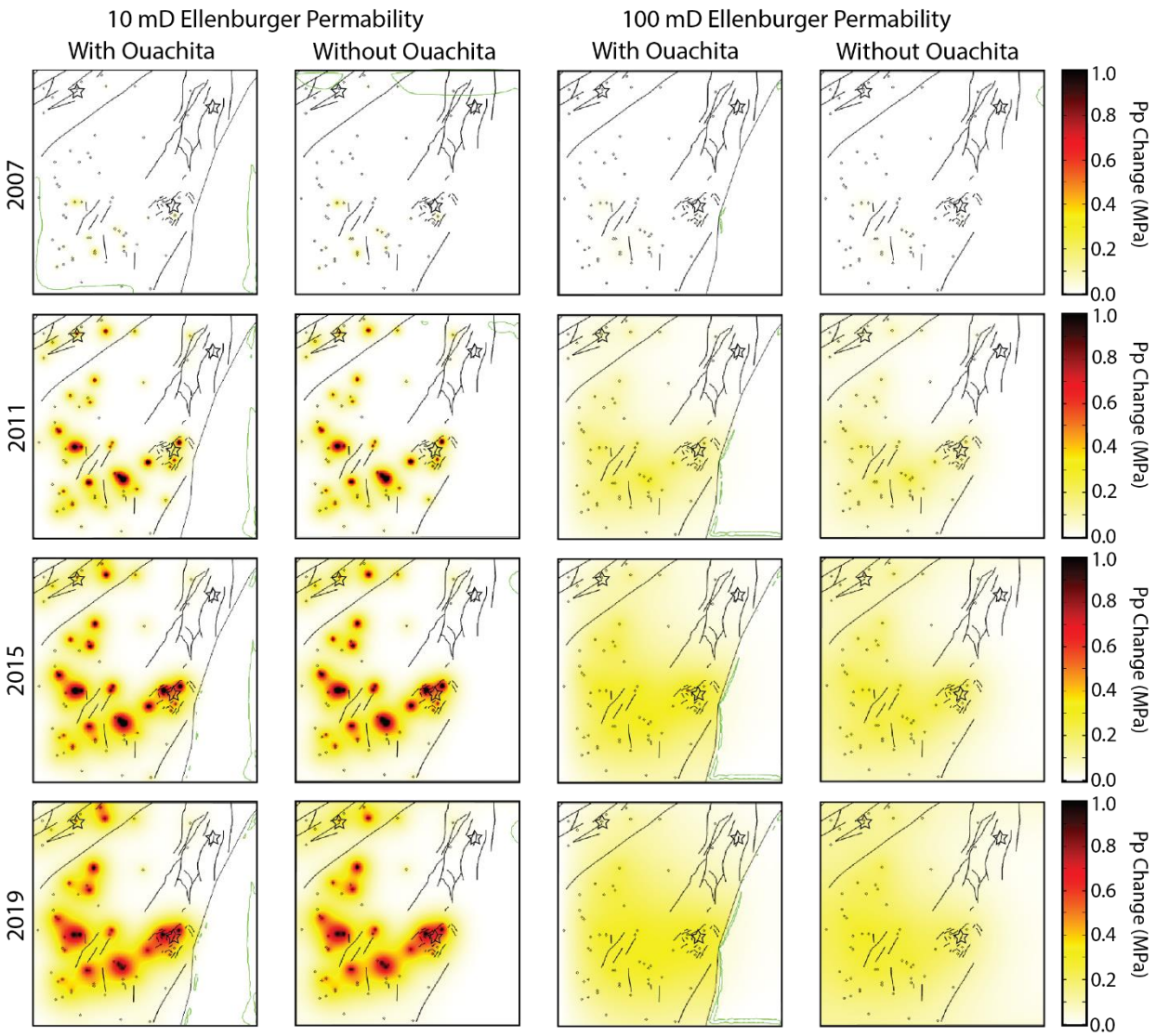

Figure 3.32. Snapshots of the modeling results from the top of the Ellenburger formation showing the pore fluid pressure changes using a variety of permeability values for the Ellenburger formation and changing how the Ouachita Thrust Front interacts with the model. Each model snapshot is taken to be from January of the labeled year. The sites of the Azle-Reno (northwest), Dallas-Irving (northeast), and Venus (southeast) earthquake sequences are represented by the black stars. The green contour line represents the parts of the model space where pore fluid pressure changes were zero.

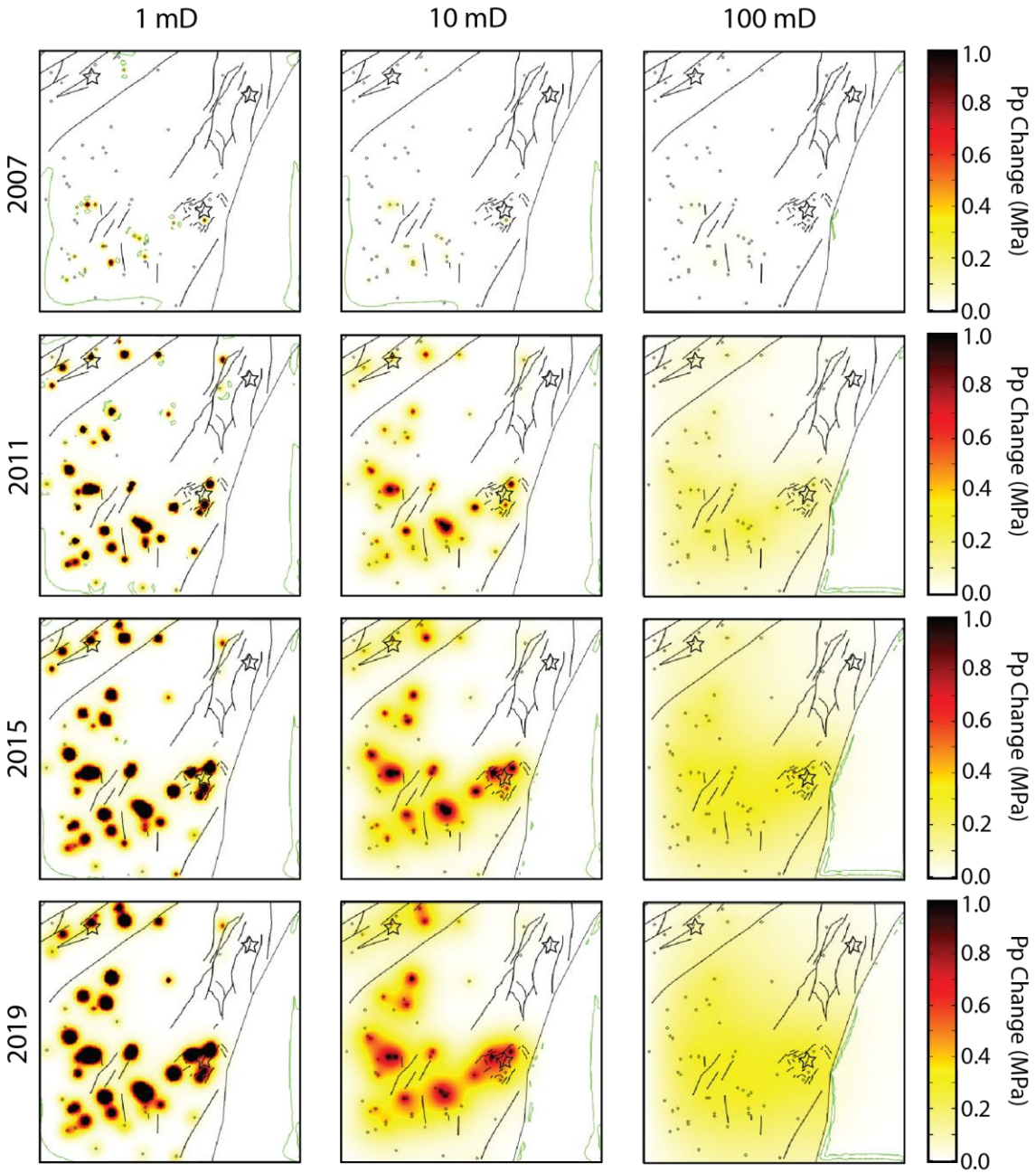

Figure 3.33. Snapshots of the modeling results from the top of the Ellenburger formation showing the pore fluid pressure changes using a variety of permeability values for the Ellenburger formation with the same boundary conditions and fault maps. Each model snapshot is taken to be from January of the labeled year. The sites of the Azle-Reno (northwest), Dallas-Irving (northeast), and Venus (southeast) earthquake sequences are represented by the black stars. The green contour line represents the parts of the model space where pore fluid pressure changes were zero.

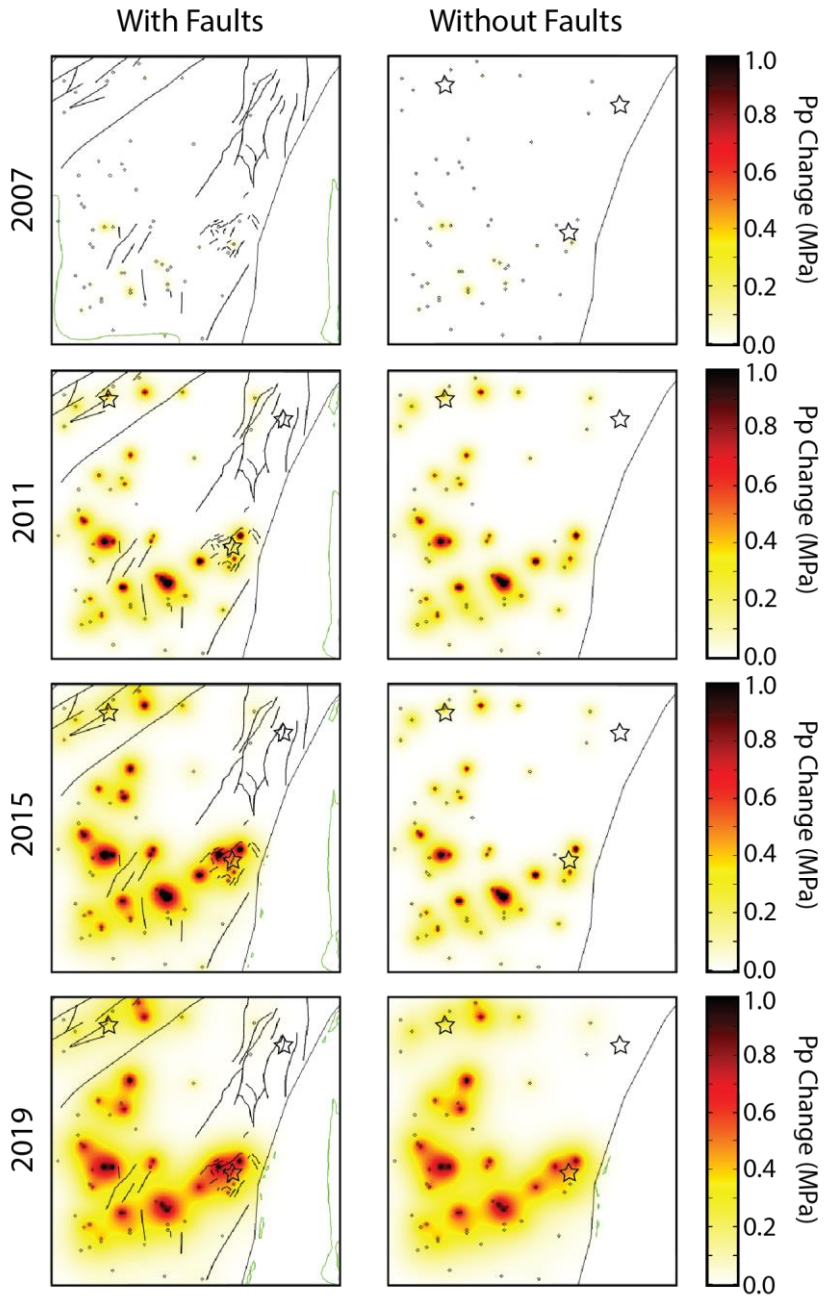

Figure 3.34. Snapshots of the modeling results from the top of the Ellenburger formation showing the pore fluid pressure changes using different fault maps for the Fort Worth Basin with the same boundary conditions and permeability values for the Ellenburger (10 mD). Each model snapshot is taken to be from January of the labeled year. The sites of the Azle-Reno (northwest), Dallas-Irving (northeast), and Venus (southeast) earthquake sequences are represented by the black stars. The green contour line represents the parts of the model space where pore fluid pressure changes were zero.

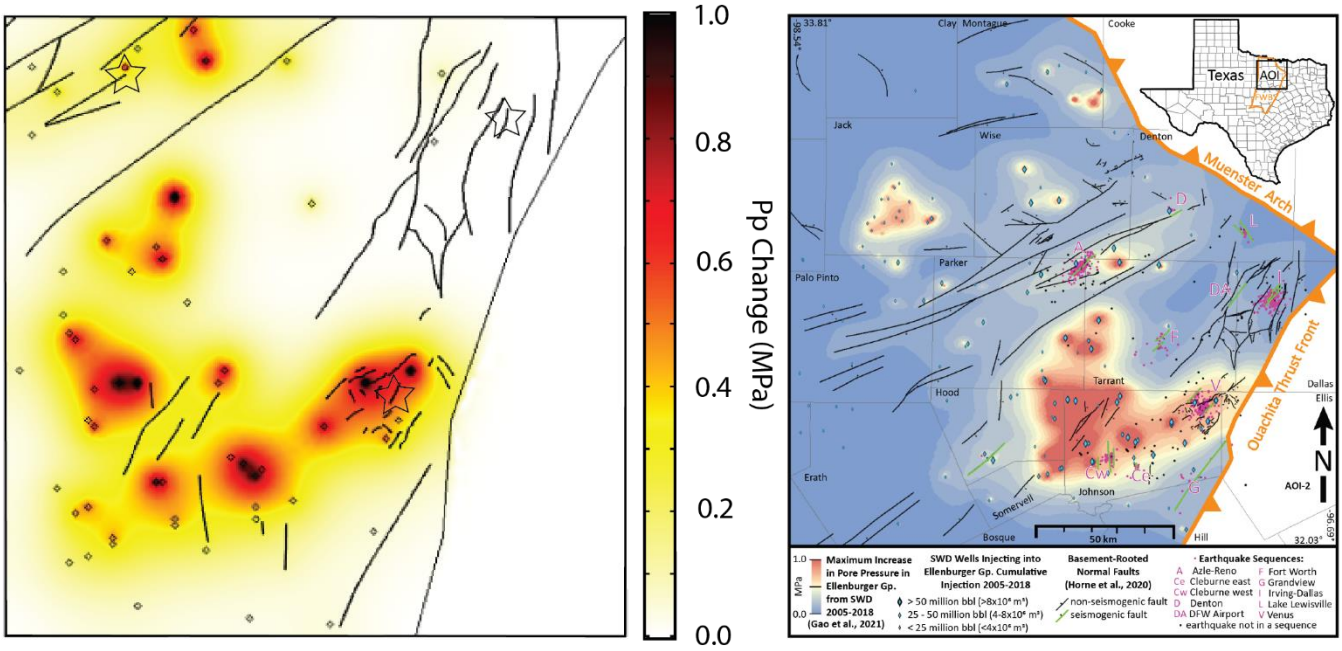

Figure 3.35. Comparison of the final modeling results from the created COMSOL coupled geomechanical model (left) and the pure fluid flow model created by Gao et al., 2021 (right). The presented COMSOL model uses a 10 mD Ellenburger permeability value, treats the Ouachita Thrust Front as an impermeable eastern boundary of the basin, and uses the fault maps presented in Hennings et al., 2019. The Gao model results were taken from a figure published in Hennings et al., 2021.

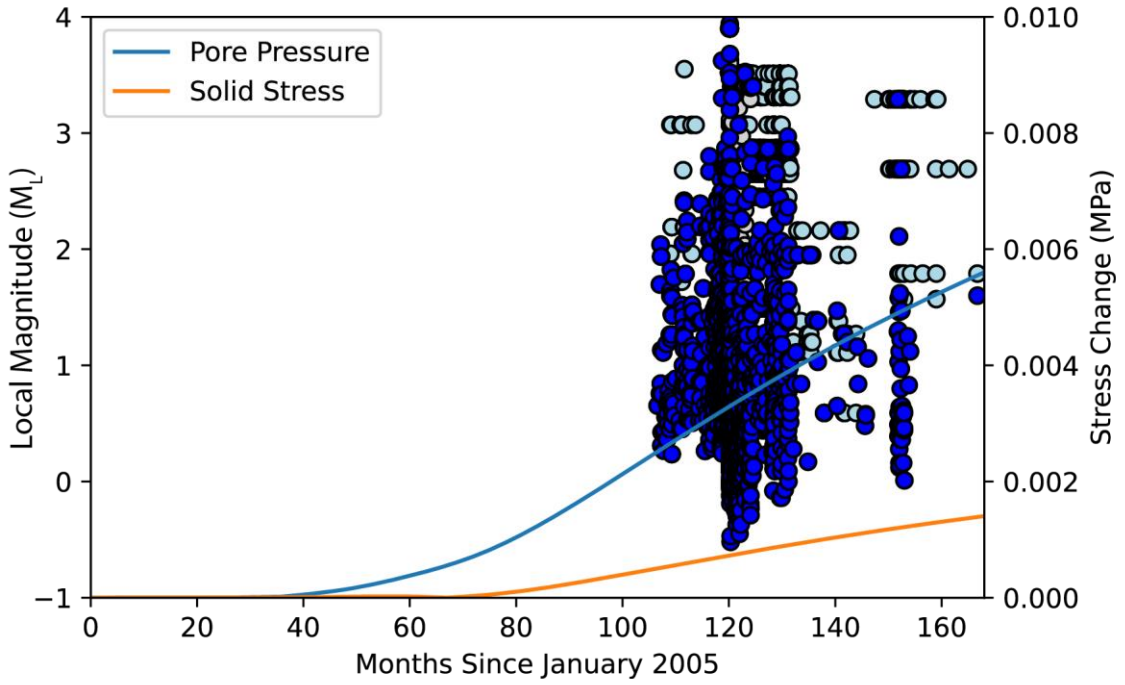

Figure 3.36. Plot of the pore fluid pressure (blue line) and poroelastic (orange line) stress changes observed at the location of the Dallas-Irving fault at the depth of the Ellenburger-Basement contact point (4 km). Also shown are the occurrence of earthquakes associated with the Dallas-Irving fault from the NTXES earthquake catalog and the occurrence of earthquakes within the Dallas-Irving (light blue) and M. Brudzinski (gray) template matching catalogs.

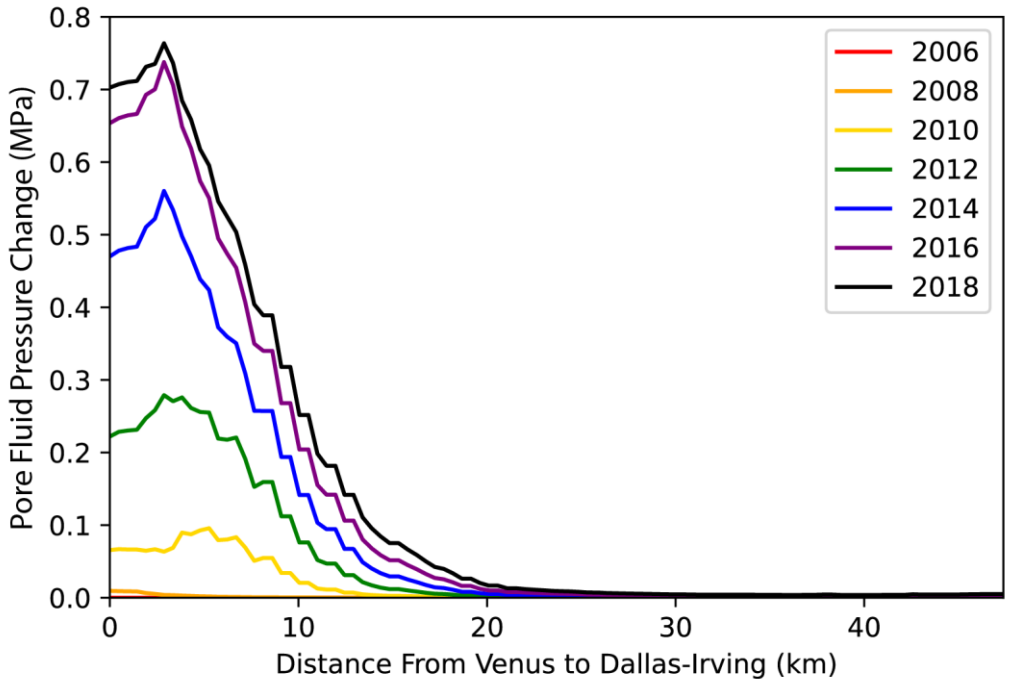

Figure 3.37. Plot showing the pore fluid pressure changes at a depth of 4 km from the model run shown in Figure 3.35 from the Venus to Dallas-Irving earthquake sites.

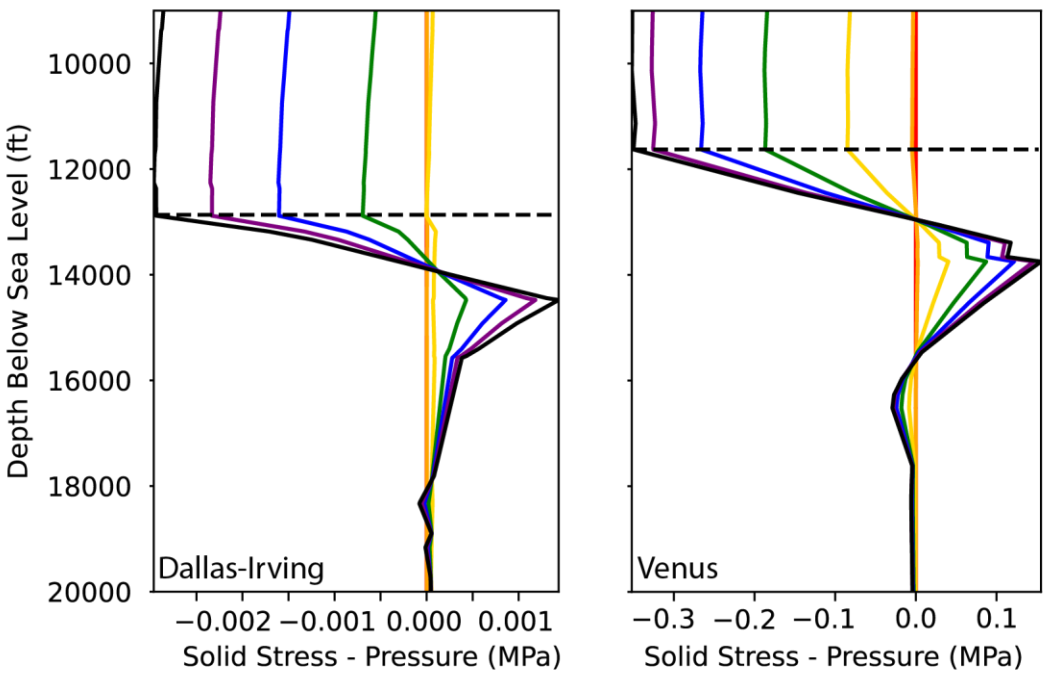

Figure 3.38. Plots showing depth profile views of the poroelastic stress changes observed at the sites of the Dallas-Irving and Venus earthquake sequences over time using the same color scheme as Figure 3.35. The dashed line represents the cutoff depth between the Ellenburger and crystalline basement formations at each sequence site.

	Cross-Correlation Data				Catalog Data					
NITER	**WT-CCP**	**WT-CCS**	**WR-CC**	**WD-CC**	**WT-CTP**	**WT-CTS**	**WR-CT**	**WD-CT**	**WT-CD**	**DAMP**
2	0.01	0.01	-9	-9	0.10	0.10	-9	-9	10	200
2	0.01	0.01	-9	-9	0.10	0.10	6	4	10	200
2	0.01	0.01	-9	-9	1.00	1.00	6	4	0.10	300
3	1.00	1.00	6	2	0.10	0.10	6	4	0.10	300
3	1.00	1.00	6	0.5	0.10	0.10	6	2	0.10	200

Table 3.1. Iteration parameters used within the *tomoDD* double-difference earthquake relocation codes which use the following parameters. NITER: number of iterations to be run using the defined weights. WTCCP, WTCCS: weighting values for the cross-correlation P and S-wave differential time values, respectively. WTCTP, WTCTS: weighting values for the catalog P and S-wave differential time values, respectively. WRCC, WRCT: residual threshold in seconds for cross-correlation and catalog differential time data, respectively. WDCC, WDCT: maximum distance in km between linked event pairs in the cross-correlation and catalog data, respectively. WTCD: relative weighting between absolute and differential time data. DAMP: dampening value used in the least squares calculations.

IT	EV	CT	CC	RMS-CT		RMS-CC		RMS-ST	DX	DY	DZ	DT	OS	AQ	CND
--	%	%	%	ms	%	ms	%	ms	m	m	m	ms	m	--	--
1	100	100	100	382	-27.7	132	-18.9	1347	200	180	226	0	197	0	43
2	100	100	100	442	15.6	142	7.4	555	64	33	124	0	145	0	44
3	100	98	100	57	-87.2	144	1.6	337	50	56	81	0	173	0	48
4	100	98	100	53	-6.0	133	-8.0	376	88	44	61	0	137	0	46
5	100	98	100	21	-0.1	82	-40.5	185	32	33	20	0	168	0	31
6	100	97	100	19	-9.1	82	0.5	187	16	15	15	0	166	0	30
7	100	97	81	23	26.0	35	-57.3	97	42	38	29	0	161	0	34
8	100	97	81	26	10.0	33	-5.6	105	21	18	17	0	159	0	33
9	100	97	80	27	3.4	33	-1.6	111	13	11	14	-17	156	0	31
10	100	95	18	26	-3.7	34	4.6	115	18	17	15	3	154	0	35
11	100	95	18	25	-2.2	32	-6.8	117	11	9	11	2	153	0	34
12	100	95	18	25	-0.5	31	-2.9	118	8	6	10	2	152	0	34

Table 3.2. Results from the 12 iterations of *tomoDD* ran using the parameters defined in Table 3.1. The columns represent – IT: Iteration number, EV: percentage of events used in each iteration, CT and CT: percentage of catalog and cross-correlation data used in each iteration, RMSCT and RMSCC: the RMS residual (in ms) and its percent change from the last iteration for the catalog and cross-correlation data, RMSST: largest RMS residual observed at a station, DX, DY, DZ, DT: average absolute value of the change in hypocenter location and origin time for each iteration, OS: shift between the cluster centroid of the initial locations and the centroid of the relocations, AQ: number of air quakes detected and discarded, CND: condition number for the system of double-difference equations.

Formation	E (GPa)	v (1)	ρ (kg/m^3)	Φ (%)	k (mD)	μ (Pa*s)	χ (1/Pa)
Upper Strata	10	0.25	2100	15	10,000	-	-
Barnett Shale	20	0.25	2500	10	0.01	-	-
Ellenburger	35	0.25	2700	4	10	-	-
Basement	50	0.25	2750	1	0.001	-	-
Faults	-	-	-	-	100	-	-
Fluid	-	-	1000	-	-	8.9e-4	4e-10

Table 3.3. Table of the major geologic and hydrologic parameters used for the defined formations within the created geomechanical model. The major parameters are as follows. *E*: the Young's modulus. *v*: the Poisson's ratio. ρ: density. Φ: porosity. *k*: permeability. μ: fluid viscosity. χ: compressibility of the fluids.

Model Realization #	Imperm-eable Ouachita?	Fault Net-work?	Ellenburger k?	Peak P_f Value	Peak P_f at Dallas-Irving	Peak P_e Stress	Peak P_e at Dallas-Irving
1	Yes	BEG	10 mD	1.338 MPa	0.006 MPa	0.969 MPa	0.003 MPa
2	No	BEG	10 mD	1.362 MPa	0.006 MPa	0.961 MPa	0.002 MPa
3	Yes	BEG	100 mD	0.369 MPa	0.046 MPa	0.118 MPa	0.011 MPa
4	No	BEG	100 mD	0.325 MPa	0.029 MPa	0.172 MPa	0.015 MPa
5	Yes	BEG	1 mD	8.341 MPa	<0.001 MPa	6.869 MPa	<0.001 MPa
6	Yes	None	10 mD	1.051 MPa	0.003 MPa	0.774 MPa	0.002 MPa

Table 3.4. Table of the model realization results using a variety of treatments of the Ouachita Thrust Front, incorporated fault networks, and permeability values for the Ellenburger formation. The modeling results focus on the pore fluid pressure (P_f) and poroelastic (P_e) stress changes observed throughout the modeling space and at the site of the Dallas Irving fault.

4. LOCAL EARTHQUAKE CATALOGS CREATED USING BROADBAND AND LARGE-N SEISMIC ARRAYS FOR THE COMMUNITY WAVEFIELDS DEMONSTRATION EXPERIMENT

4.1. Introduction

The IRIS Community Wavefields Experiment recorded local and regional earthquakes, teleseisms, infrasound, regional distance-controlled sources and the 3 September 2016 M5.8 Pawnee, OK, earthquake. The experiment design focused on recording of the full wavefield using 3-component nodal technology supplemented by infrasound and broadband seismic sensors (Figure 4.1; Sweet et al., 2018). The unique design included a gradiometer array, Golay array, infrasound array, and any number of sub-arrays built from the nodal stations deployed along East-West and two North-South roads. Siting in north-central Oklahoma allowed the community experiment to take advantage of high seismicity rates during the 2016 period (Sweet et al., 2018; Keranen et al., 2014; Schoenball and Ellsworth, 2017; Zhai et al., 2019; Skoumal et al., 2020). Prior to deployment there were at minimum two known active faults generating small (magnitude<3) earthquakes directly below the footprint of the experiment. The 363 nodes were deployed for a 30-day period in June-July 2016, and the broadbands and infrasound arrays were left in the field through November 2016, in part to capture the aftershock sequence of the Pawnee earthquake to the east.

The CWDE scientific justification arose from a series of community-developed white papers in 2015. The goal was to provide a rich dataset that would capture the full wavefield using nodal technology and a relatively large number (large-N) of stations (>350 stations) and address open questions about the physical mechanisms and driving processes of injection induced earthquakes. The CWDE also allowed testing of unique array geometries and was designed to propel development of algorithms to efficiently ingest a mix-mode data. From a practical standpoint, the CWDE provided IRIS and the seismology community training on nodal deployments and spurred development of best practices for rapid nodal deployments and archiving. The dataset has been used in a range of studies to date (Sweet et al., 2018; Fan and McGuire 2018; Li et al., 2018; Bolawinra and Langston, 2021).

The goals of this study were two-fold:

(1) I provided the first time-continuous earthquake catalog, including focal mechanisms, for the CWDE. The catalog was built following standard approaches from a temporary seismic research network and developed with the Antelope environmental monitoring system (Boulder Real Time Technologies, brtt.com). For the initial catalog, I focused on automatically and manually detected P and S arrivals recorded by all broadband stations located in north-central Oklahoma in the June - November time frame, and only reported earthquakes within 15 km of the CWDE experiment. Broadband stations from the CWDE were combined with select stations of the XR broadband network operated by Cornell University. I also provided arrival times, local magnitudes, error estimates, and focal mechanisms and developed a revised 1D velocity model for the CWDE footprint. The catalog and associated products, hereafter called the CWDE_BB catalog, are provided to seed further studies of small magnitude, local distance, induced

earthquakes and serve as a benchmark dataset for detection, association, and location algorithm development (Appendix C.1).

(2) I provided proof of concept for development of rapid earthquake detections using nodal data that in turn could be used to develop hypocenter catalogs. I applied a local similarity detection method that takes advantage of the close station spacing of the nodes (72.8 m). Here, I focused on building a detection list, hereafter the CWDE_NDL catalog, using data collected on all channels of the three-component nodal stations. I described the relationship between the processed waveform data and the local similarity results and identified how the local similarity data can be used to identify network wide average P- and S-wave arrival times for each event detection. I also showed that the these identified network P- and S-wave arrival times can be used to discern whether a detected event was a defined local (<15 km epicentral distance to center of array) or regional (>15 km epicentral distance to center of array) event. I also compared the event detection capabilities of the local similarity technique when compared to the traditional analyst-based approach of manually identifying phase arrival times to detect events. The goal was to rapidly and efficiently use the nodal data to create a high-resolution detection list for P- and S-wave onsets that could then serve as a target for refined analyses. The CWDE_NDL catalog is not designed to replace nor be used for research in the same manner as the CWDE_BB catalog (Appendix C.2).

4.2. Methodology for the CDWE Broadband (CDWE_BB) Catalog

I used the Antelope software to develop the CWDE_BB catalog. The analysis workflow relies upon accessing the publicly available data stored within the IRIS Digital Management Center under Federation of Digital Seismic Network (FDSN) code YW 2016. The large CWDE dataset was accessed through the PH5 repository, as now standard with high sample array data

from temporary experiments. I used the 18 broadband stations deployed as part of the CWDE and 25 broadband stations deployed as part of the "Seismicity near the Nemaha fault in northern Oklahoma" (OK GAP) project (FDSN XR 2016-2017; Keranen, 2016) (Figure 4.1). For both experiments, I only analyzed the high-gain, high sample rate 3-component broadband channels sampled at 100 Hz (CWDE) or 200 Hz (OK GAP). Both experiments returned high quality continuous data streams with no sensor changes. Both experiments deployed Guralp sensors on RefTek RT-130s [CWDE: CMG-3ESP 120s sensor and OK GAP: CMG3T]. Additionally, I provide the 2016 injection volume data for the wells shown in Figure 4.1 in Appendix C.3 to provide further benchmark data sets for additional studies.

The CWDE_BB catalog was built using a standard workflow within the Antelope processing system for offline research-based temporary experiments. The automatic detection software (dbdetect) used short-term average/long-term average ratio (STA/LTA) to create a list of detection times which were then associated and located using a spatial grid-search based approach based on predicted travel-times through the global iasp911D velocity model (dbgrassoc). The STA/LTA parameters were as follows: the short term and long-term average time windows were taken to be 0.1 and 10.0 seconds respectively; and a signal-to-noise threshold value of 2.5 was required for a detection to be recorded. No initial filtering of the data was used for the broadband based STA/LTA detections. All automated arrivals and locations were manually verified, and P-wave first motions were determined without use of filters; earthquake onsets exhibited high signal-to-noise over the 10-40 Hz frequency range. The earthquakes were then relocated using the GenLoc library (dbgenloc; relocate) (Pavlis et al., 2004), which provided hypocenter error estimates based on the 68% confidence range ellipses derived directly from the covariance matrix of each event and an error bound based on velocity model uncertainty.

Because I was concentrating on events within 20 km of the CWDE footprint, I revised the Oklahoma Geological Survey (OGS) 1D velocity model used for the state-wide earthquake catalog using OGS stratigraphic reports (Crain and Chang, 2018a, 2018b; Figure 4.2). Similarly, the local magnitude calculator was modified to reflect the distance attenuation used by the OGS (Walter et al., 2020). It is necessary to point out that this procedure was first applied to CWDE stations for reasons discussed further below.

The CDWE_BB catalog can temporally be split into two periods. The first period encompasses the time when both broadband and nodal stations were active as part of the CWDE array (06/25/16 – 07/19/20), and the second section represents the time period where only broadband CWDE stations were active (07/19/20 – 11/14/16). The more spatially and temporally extensive OK GAP experiment spans the entire CWDE experiment time period. The June/July component of the CWDE_BB catalog is the benchmark dataset for comparison with nodal-based analysis techniques. For this comparison time period only, analysts visually reviewed all continuous waveform data collected by the CWDE array broadband stations and added hundreds of earthquakes missed by the automated procedure. For an analyst to record an event within the CWDE_BB catalog it was required that they observe both P- and S-wave phase arrivals on a majority of the CWDE broadband stations (>8 stations, minimum 16 associated phases) using unfiltered waveform data. This was done to construct the best dataset for comparison with automated procedures using the nodal array.

Southern Methodist University (SMU) first developed a version of the CWDE_BB catalog in 2017 for informal rapid distribution to those interested in working on the community dataset. That catalog, however, had a fundamental issue - namely, the footprint of the CWDE project was small but the seismicity rate around the experiment was high (Figure 4.3). There

were hundreds of earthquakes well-recorded on the CWDE experiment, but located outside of the broadband array, such that there was no station coverage beyond a few kilometers to minimize origin time/depth tradeoffs. Studies of induced earthquakes related to injection rely on low-uncertainty depths and origin times to correctly identify the time and space relationships to local anthropogenic activity (Ellsworth, 2013; Keranen et al., 2013; Frohlich et al., 2016; Pei et al., 2018; Cochran et al., 2020). This issue can best be visualized by looking at the greatest azimuthal P-wave separation (GAP) of this circa-2017 catalog (Figure 4.4). Poor azimuthal coverage also severely limited focal mechanism quality. Inclusion of the OK GAP stations, publicly released in 2019, resolved this issue. Collaborators at SMU manually added P- and S-wave onset times and first motions for OK GAP stations located within 20 km of the CWDE and for earthquakes of magnitude >=1.5. This improved CDWE_BB catalog is reported in this chapter.

4.3. Focal Mechanism Determination

The focal mechanisms associated with the CWDE_BB catalog incorporated both P-wave first motion and S to P-wave amplitude ratios (S/P) to constrain the solutions (HASH; Hardebeck & Shearer, 2002, 2003). The P-wave first motion and S/P amplitude ratio data was collected from the CWDE array and surrounding XR network broadband stations to minimize azimuthal coverage deficiencies associated with events located outside of the aperture of the CWDE array (Figure 4.4). I generated focal mechanisms only for events with magnitudes exceeding M_L1.5 located using a minimum of six broadband stations with associated P and S-wave arrivals and exhibiting signal-to-noise ratios greater than three. Calculation of the S/P amplitude ratios follows the procedures outlined in Hardebeck and Shearer (2003). Namely, as the catalog consists of local earthquakes located using primarily upgoing ray paths that are largely

constrained to the shallow structures in the region, both the site and path effects can be corrected using the same station correction value. This station correction value was calculated by examining the full S/P amplitude ratio distribution of the combined CWDE and XR broadband station arrays and corrects each individual station amplitude ratio values under the assumption that the full array provides full azimuthal coverage for the events. Each mechanism is assigned a quality rating which corresponds to the RMS fault plane uncertainty with ratings such as: RMS uncertainty ≤25° (A), RMS uncertainty ≤35° (B), and RMS uncertainty ≤45° (C). In total, 124 A and B quality focal mechanisms were generated which consisted almost entirely of predominantly strike-slip solutions (95%) with few normal (4%) or reverse (1%) focal mechanism solutions according to standard mechanism definition criteria (Kagan, 2005; Álvarez-Gómez, 2019; Figure 4.5). The mechanism nodal plane solutions have high agreement with the recorded regional stress field in northern Oklahoma and with the fault lineaments imaged by the earthquake locations (Figure 4.3).

4.4. Methodology for the Nodal Wavefields Catalog

The CWDE_NDL is meant to explore how an analyst might be able to begin to create an earthquake catalog by examining data collected rapidly and efficiently on large-N style arrays. The methodologies used to create the benchmark CWDE_BB catalog (06/25/16 – 07/19/16) required the use of several analysts and numerous weeks of labor time to manually review all waveform data for that time period on the combined YW and XR network broadband stations. This manual labor required for the benchmark dataset is not feasible using the 363-station nodal array and thus a different methodology must be employed to more efficiently analyze the large quantity of data associated with this large-N style array (~1.76 TB of data). Here, I employed the

local similarity detection methodology for high-resolution seismic event detection to create a list of event detection times using the CWDE nodal array dataset (Li et al., 2018).

The local similarity event detection method is based on cross-correlating waveforms of proximal stations in order to increase recorded signal-to-noise ratios of waveforms and to exploit coherent properties of the wavefield, as outlined in Li et al. (2018). This type of waveform stacking is predicated on the assumption that noise signals recorded on stations near one another (<500 m station spacing) will be less coherent than earthquake signals detected across those same stations. This method defines the "local similarity" recorded on a station at a given point in time to be the signal correlation between the station itself and its nearest neighbor stations. A key assumption of this method is that the signals recorded by two separate but closely spaced stations should have nearly identical ray paths when generated from the same source. A similar assumption is made when examining double-difference travel time values for the purposes of earthquake relocation (see Chapter 3). The very closely spaced station spacing needed for the local similarity technique is unlikely to hold for arrays with uneven station spacing with larger inter-station distances or small number of stations. Here, the CWDE nodal array average station spacing of 72.8 m is densely spaced for use in the local similarity methodology.

The local similarity event detection algorithm correlates and stacks similar waveforms taken from close distance stations to increase signal and reduce noise correlation. The four processing steps include: selecting the nearest neighbors to each central station, calculating the local correlation using the central station and its nearest neighbors, stacking the resulting 363 local correlation values to generate a network correlation value, and lastly using a correlation threshold detector to determine event detection times (Figure 4.6). The first step involves using a nearest neighbor search algorithm, which treats each of the 363 nodal stations in the array as its

own central station. In this study I used the four nearest neighbor stations to each central station to calculate the local correlation values; the average nearest neighbor distance across the array is 72.8 m and ranges from 0.0 m to 194.3 m (Figure 4.7).

The local correlation of each central station is taken to be the average of its sliding-window normalized cross-correlation values. I allowed the sliding time window to shift within the lag time of the cross-correlations with its nearest neighbors to account for any arrival time differences between stations. During this process I first calculated each central station local correlation value independently as if each local cluster of a central station and its nearest neighbors were a separate individual network of stations. Li et al. (2018) defined the peak cross correlation function between a central station and its nearest neighbors as:

$$s_{ij}(t) = \frac{\left|\sum_{m=-M}^{M} u_i(t+m\delta)u_{ij}(t+m\delta+l\delta)\right|}{\sqrt{\sum_{m=-M}^{M} u_i^2(t+m\delta)\sum_{m=-M}^{M} u_{ij}^2(t+m\delta+l\delta)}} \quad (4.1)$$

where u_i is the signal recorded on the i^{th} master station and u_{ij} is the signal recorded on the j^{th} neighbor station, δ is the sampling rate, $(2M+1)\delta$ is the sliding window length, and $L\delta$ is the maximum time lag allowed (determined by a limit on the wave slowness and distance between the central and neighbor stations). After the local correlation of all 363 central stations has been calculated for a given time window, then the resulting cross-correlation values are averaged into the new local similarity metric that describes the correlation of each master station. The overall network local similarity is calculated by directly stacking the resulting local similarity traces of each master trace (Figure 4.6), which allows the technique to be used as a general detection method without the need to assume any specific wave type or velocity model (Li et al., 2018).

To determine the start and end times for an event detection we set threshold values for the overall network local similarity value. These detection threshold values are calculated over a sliding time window with a length of five minutes across the entirety of the full local similarity data (Figure 4.8). A five-minute time window was chosen for the CWDE data set as this window accounted for longer temporal trends in the overall network local similarity values while remaining unbiased during periods of high detection activity that caused smaller time window on threshold values to become too high. I calculated the median absolute deviation values of the network local similarity data within each time window and then set the "on" threshold value signifying the start of an event detection to be ten times the median absolute deviation value. Next, the end time of an event is determined once the local similarity value falls below the "off" threshold value which is set at five times the median absolute deviation value within the same time window. The CWDE_NDL detection list contains 3467 event detections for the time period of the nodal array (06/25/16-07/19/16).

The detections in the CWDE_NDL detection list each contain an associated detection time corresponds to the time when the "on" threshold value was surpassed and a time value corresponding to when the "off" threshold value was subsequently reached. However, this "on" detection time value, hereafter referred to as the LS detection time, is not the same as the peak local similarity value, hereafter referred to as the LS peak time, associated with an event detection (Figure 4.9). To understand the physical meaning of the local similarity results with regards to the typical origin and first-arrival onset times (P or S) associated with an event detection I examined a set of local similarity detected events that were also recorded within the CWDE_BB catalog (Figure 4.6). First, I compared the local similarity results from each of the three nodal station channels and observed that these events were detected on each channel

(Figure 4.10). Next, I compared the recorded LS detection times across the three channels and found that while there were slight variations in the actual recorded time value, the variability in LS detection times was small (<0.5 s). The first notable feature of the local similarity results is that none of the LS detection times recorded on the three channels correspond to the origin time of an event (Figure 4.11). Instead, I observed that the LS detection times best matched the arrival onset time of the first P-wave to have reached the nearest broadband station to each event; specifically, it is the LS detection time recoded by the vertical channel best match P-wave arrival onset time. The LS detection times recorded by the horizontal channels do occasionally match the broadband recorded first P-wave onset times, but there is a larger amount of variability when compared to the vertical channel LS results and so here are treated as secondary measures of possible P-wave onset times. However, the LS peak times recorded on the horizontal channels match well with the S-wave arrival onset times on the broadband stations (Figure 4.11). This relationship most likely reflects the fact the periods of peak signal coherency across the horizontal channels of the nodal network best match those periods of the high amplitude S-waves travelling across the network. Thus, I define the LS detection times from the vertical channel results and the LS peak times from the horizontal channel results to represent the first P-wave and S-wave arrival times across the nodal network, respectively. As a reminder, the goal is to explore ways to build a rapid catalog of thousands of earthquakes that would then need to be refined and quality controlled.

The local similarity results allow for an estimate of the P- and S-wave arrival onset times for the entirety of the nodal network (not at each nodal station), but it becomes possible to use the resulting S-minus-P-wave relationship to estimate distance and determine if the earthquake is a local or regional event. For this study, a local event is defined to have occurred within 20 km of

the defined center point of the CWDE array and a regional event is any event which occurred >20 km away from the same point. The relationship between the station-to-event distance and the S-minus-P time associated with that event is well understood (Kennett and Engdahl, 1991) and here I estimated, using the 1D velocity model I created for the CWDE study area, that the S-minus-P time for an event located 20 km away from the center point of the array is 2.3 s. That means that all CWDE_NDL detected events whose S-minus-P time is less than 2.3 s are treated as local events and can thus be directly compared to the CWDE_BB catalog of local seismicity recorded by the broadband stations. It should be noted that because the local similarity technique does not calculate a proxy P- and S-wave arrival time for each individual nodal station, the P- and S-wave onset times associated with each event in the CWDE_NDL at best represent an estimate. While it is possible to discern whether a local similarity detected event is a local or regional event, to estimate a proper event location further phase identification on additional stations would be required. Here, the CWDE_NDL includes: the LS detection time recorded on the vertical channel, the average LS peak time recorded from each of the horizontal channels, and the latest "off" detection time associated with the horizontal channels (Appendix C.2).

4.5. Earthquake Spatial and Depth Distributions

The events in the CWDE_BB catalog represents the local seismicity detected around the CWDE array (<20 km epicentral distance to stations). The earthquakes spatially cluster into a series of linear features interpreted to be faults active during the period of deployment (Figure 4.3). Most seismicity occurs on faults that had prior records of seismicity (e.g., Schoenball & Ellsworth, 2017); however, there are some clusters of seismicity northwest of the array that likely indicate previously inactive and/or unmapped faults. The majority (99.2%) of earthquakes within the CWDE_BB catalog are located within the crystalline basement formation at depths

typically ranging from 2-8 km (Figure 4.12). Given the Arbuckle-basement boundary depth is 1.8 km below sea level here and the average depth uncertainty is 492.5 m, we suggest that some of the few (0.8%) events located above the Arbuckle-basement boundary are likely located within the basement. The CWDE_BB earthquake depth values are deeper than previous depth estimates in this region. However, previous catalogs of seismicity (e.g., Schoenball and Ellsworth, 2017; Walter et al., 2020) were reliant on data collected at regional distances without a local station to constrain depth values and did not have the available formation top data to construct a more localized velocity model to locate the events, as I have done here. This depth distribution for events occurring along basement-seeded faults that underlie a major fluid injection formation (Arbuckle formation) is expected given the working hypothesis that the seismicity has been induced by nearby fluid injection activities.

4.6. Fault and Stress Characterization

The CWDE_BB earthquake locations show several primarily ENE-WSW trending faults throughout the study area (Figure 4.3). These faults are mainly shorter length (<5 km) basement-seeded faults with depths ranging between 1.8 and 9.0 km. The focal mechanisms indicate nearly vertical oblique strike-slip faults with strikes ranging from 52-107°. Activation of near-critically stressed faults of these strikes would be expected given that the regional stress field in the study area is a strike-slip regime where the maximum horizontal stress is approximately 72° (Alt and Zoback, 2017). There are some higher quality (A and B grade) mechanisms that describe normal or reverse faulting; these solutions are not characteristic of any individual fault in the study area and may be associated with earthquakes occurring off the main fault structure and within the damage zone. There are clusters of earthquake locations throughout the study area which, either due to a lower number of events within the cluster or due to higher location uncertainties, do not

appear to have occurred on any previously identified fault. However, the majority of seismicity in the CWDE_BB catalog does occur along previously or newly identified faults consistent with activation within the regional stress field of northern Oklahoma (Schoenball and Ellsworth, 2017).

4.7. Magnitude Distribution

The CWDE_BB catalog contains lower magnitude (M_L<2.0) events, with the largest magnitude event being a M_L3.8 event on 10/21/2016 (Figure 4.12). As previously discussed, the CWDE_BB catalog can temporally be split into two distinct time periods: June 25th – July 19th and July 19th – November 13th, 2016. The first period corresponds to when the nodal stations were still active and recording seismicity and thus the local similarity technique was able to be used to detect events. Thus, during that period analysts were tasked with recording every earthquake possible to create the most complete catalog of seismicity to serve as a benchmark for comparison purposes. During the latter period when only the broadband stations were active, only events detected using the previously discussed STA/LTA detector were incorporated into the catalog. Due to the severe difference in the detection techniques applied to both periods, I chose to analyze the magnitude distributions of both periods separately. Specifically, I calculated separate magnitude of completes (Mc) and b-values for each period using the 90% goodness-of-fit methodology (Wiemer and Wyss, 2000; Figure 4.13). What I found is that during both periods the b-values reported for the CWDE_BB catalog are below 1.0, although the b-value associated with the first period is higher (0.757 ± 0.03 vs. 0.628 ± 0.02). Additionally, the Mc values associated with the first period where the analysts were tasked with picking as many events as possible regardless of the low magnitudes was far lower than the period where only STA/LTA detected events were incorporated (0.3 vs. 1.1). The low b-values associated with each period of

the catalog may be reflective of the strict distance cut-off applied to the catalog where many smaller magnitude events which may have been located outside of the 20 km allowed radius around the array could have increased the overall b-values associated with the first period. The low b-value of the second period is more likely due to the same issues of events occurring outside the allowed radius of included events and the STA/LTA detector simply missing many low magnitude events which an analyst could have detected (a problem noted in Chapter 2 for north Texas as well). When I examined the temporal evolution of the Mc and b-values of the CWDE_BB catalog I found that during the earlier period of time the b-value was close to 1.0, however when larger (M_L>2.0) events began to occur the b-value quickly lowered and remained steady at its near 0.7 value (Figure 4.14). Ultimately, I do not believe that the lower b-values reported for the CWDE_BB catalog reflect interpretable changes to the regional stress field but rather are associated with the overall completeness of the catalog; b-value should not be over interpreted.

4.8. Comparison of Catalog Detection Capabilities

Next, I compared the earthquake catalog created using traditional detection-association-location techniques to the event detection list created using large-N based techniques. Overall, during the time period that the CWDE nodal stations were active (06/25/16 - 07/19/16), there were 3467 events detected using the local similarity technique and 1084 events detected and located using traditional approaches (Figure 4.15). I calculated a cross-correlation value of 0.71 between the daily event detection rates of the entire CWDE_NDL detection list and CWDE_BB catalog. I hypothesized that this moderate correlation between the two methods reflected the fact that the local similarity technique is capturing both analyst-missed local distance earthquakes and regional events associated with larger magnitude events across Oklahoma. Analysts removed

regional earthquakes from the CWDE_BB. When comparing the event detection rates of the local detections within the CWDE_NDL detection list using S-minus-P times and the CWDE_BB catalog I found a cross-correlation value of 0.77. This higher value possibly represents the filling in of missed events associated with those periods of higher seismic activity recorded within the CWDE_BB catalog. Using the CWDE_BB catalog as my benchmark catalog, I observed that the nodal based local similarity technique was able to detect 97.0% (1052 out of 1084 detections) of the events located by analysts using the broadband data. For events shared between the catalog and detection list, I additionally cross-reference all P- and S-wave arrival times from the CWDE_BB catalog against the estimated P- and S-wave onset times from the local CWDE_NDL detection list. When examining the raw waveform data and associated local similarity results from the time windows around the 32 missed detections, I found that there was enough signal coherency that I as a human analyst would have chosen to identify an event at those times (Figure 4.16). However, the reason that these events were not included in the CWDE_NDL detection list was due to "on" and "off" value detection threshold values used for the local similarity analysis. Attempts to re-tune these threshold values in order to incorporate these 32 missed detections into the full CWDE_NDL detection list led to significant numbers of false detections, however. The parameters for the local similarity technique described within the Methodology section were selected based on analysis of sample subsets of the CWDE data set and were tuned to limit the number of false detections within the CWDE_NDL data set. It is possible that lowering the event detection threshold values in the local similarity analysis would allow me to detect the missing CWDE_BB events, however it would also introduce many detections that would be classified as false detections. A comparison of the events detected by both techniques shows that the missed events were of typically smaller

magnitudes (only 1 event >M_L0.5) with the average magnitude of a non-detected event being M_L-0.06 (Figure 4.17). In Figure 4.17, I plot the associated CWDE_BB and CWDE_NDL detections with their distance relative to the central point of the array, which is defined to be the average latitude and longitude location of the combined YW broadband and nodal stations (Figure 4.1). When examining the spatial and depth distributions of the non-detected events I found no strong correlations between the depth or lateral distance from the central point of the nodal array and the detection capabilities of the local similarity technique. There does appear to be some correlation between the magnitude of the events in the CWDE_BB catalog and the ability of the local similarity technique to detect said event. Namely, the local similarity technique detects 99.6%, 98.2%, and 97.0% of CWDE_BB events with magnitudes greater than M_L0.5, M_L0.0, and M_L-0.5, respectively.

There are 2415 event detections within the CWDE_NDL detection list that are not associated with any event within the CWDE_BB catalog. While the local similarity detection technique did miss some lower magnitude events included in the CWDE_BB catalog, it does still significantly increase the signal-to-noise ratio of many small magnitude events, allowing them to cross the imposed local correlation thresholds (Li et al., 2018). Additionally, as stated earlier, the CWDE_BB catalog consists of only local events in the surrounding area around the CWDE array and so excludes regional distance events that would still have been detectable by an analyst. As the seismicity rate in Oklahoma was very high during the period the CWDE array was active (Zhai et al., 2019; Skoumal et al., 2020) we know that there were numerous regional events detected using local similarity but not included in the CWDE_BB catalog. Here, I initially utilized the local similarity methodology as described in Li et al. (2018) which focuses on the detection of events and does not directly allow for event location estimates to be made. By

expanding the local similarity analysis to incorporate all three channels from the nodal stations and examining the relationship between onset detection and peak local similarity values I was able to estimate network-wide first P- and S-wave arrival times for the CWDE_NDL detections. I was then able to use the estimated S-minus-P wave arrival time relationship using the 1D velocity model and recorded first P- and S-wave arrival times from the CWDE_BB catalog events to separate out the local and regional CWDE_NDL detections. The results from this S-minus-P time analysis of the CWDE_NDL detection list were that there were 2035 and 1432 associated local and regional events detected within the detection list, respectively. When examining the events detected using the local similarity technique that were missed by the human analysts, I noted that the majority of missed detections occurred during those period of high seismic activity as reported in the CWDE_BB catalog (Figure 4.17). This pattern of an automated modern event detection technique yielding more event detections but with a similar temporal pattern of seismicity rate to traditional event detection techniques was also observed in my template matching analysis of the Dallas-Irving earthquake sequence in the FWB (Chapter 3). While this strong temporal correlation between the detection patterns of both techniques could be interpretated as the local similarity technique making false detections, the local similarity results when compared to the raw waveform data even for those events not in the CWDE_BB catalog do show measures of signal coherency that lead me to interpret those detections as true events (Figure 4.18).

The total data volume associated with the 18 CWDE broadband stations is approximately 11.6 GB while the 363 CWDE nodal stations is approximately 1.76 TB making manual review of the nodal data unfeasible. However, when utilizing the local similarity codes using the high-performance computing center ManeFrame II at SMU the analysis of the full nodal dataset was

completed in under 48 hours. The computational resources used during this study were 16 250 GB memory nodes with 36 cores each. In comparison, the traditional combination of automated and manual analysis of the CWDE broadband data to detect and locate as many events as possible during the time period that the nodal stations took many months. This significant difference in time required to analyze the datasets exemplifies how automatic detection techniques such as local similarity can be used as rapid catalog creation tools when dealing with large-N style arrays for more targeted analysis. However, the weaknesses of the local similarity technique, namely its inability to locate an event and calculate a magnitude, also show the importance of having a strong benchmark catalog to compare to and thus better understand said results.

4.9. Discussion

The CWDE dataset and the associated CWDE_BB catalog presented here represent an opportunity for researchers to test and examine a multitude of broadband and large-N based approaches to the analysis of seismic data in a region of high seismic activity. Here, I utilized the local similarity technique to create a rapid detection list using the large-N style nodal network which contained nearly 3500 event detections over a period of approximately 25 days. However, due to the limitations of the local similarity technique I would need to incorporate other methods of analysis such as automated phase pickers (Chen and Holland, 2016; Ross et al., 2018; Zhu and Beroza, 2019; Walter et al., 2021) or back-projection techniques (Ishii et al., 2005; Hansen and Schmandt, 2015; Kiser and Ishii, 2017) to locate these events. In either case, the created catalog of events would represent an entirely automated list of event detections and locations which could be biased due to limitations associated with the techniques or errors in the chosen parameters used within the techniques. To better assess the accuracy or precision of these

catalogs it is important to have a proper benchmark. It is in these cases that the CWDE_BB catalog can best be utilized. Due to the stringent cataloging of events during the period the nodal network was active, the CWDE_BB catalog can be considered as the true benchmark for the detection capabilities of an analyst using the CWDE dataset. Therefore, when creating a catalog focused on using automated detection and location (Grigoli et al., 2013; Hansen and Schmandt, 2015; Perol et al., 2018) techniques one should look to the CWDE_BB catalog as a starting point for assessing the validity of the resulting event catalogs. Additionally, the CWDE_BB catalog can serve as an important training catalog for the purposes template matching (Skoumal et al., 2014, 2020; Scales et al., 2017; Ross et al., 2019) and machine learning (Perol et al., 2018; Li et al., 2018b; Kong et al., 2019; Zhu and Beroza, 2019; Walter et al., 2021) approaches for creating catalogs of seismicity. Template approaches rely upon the availability of previous characteristic recordings of seismicity related to the regions of study that can often be limited to a few moderate to larger magnitude events. However, with the CWDE_BB catalog any such template matching approach now has access to a catalog of over 1000 events with magnitudes ranging to even below M_L0.0.

The seismicity recorded within the CWDE_BB catalog describes several primarily E-W trending strike-slip faults with earthquakes located within the crystalline basement (Figure 4.3). The depth distribution and prior history of seismicity in northern Oklahoma suggests that like other sites of induced seismicity occurring within a basement formation underlying a major fluid disposal formation (Ellsworth, 2013; Keranen et al., 2013; Scales et al., 2017; Yoon et al., 2017; Rubinstein et al., 2018; Quinones et al., 2019) that injection associated stress changes are the main driving force of seismicity in the study area. Previous regional monitoring-based catalogs of seismicity for this portion of Oklahoma described seismicity along some of the faults imaged

by the CWDE_BB catalog, but due to limitations in location capabilities placed most of these earthquakes at much shallower depths (Schoenball and Ellsworth, 2017). Utilizing the local YW and XR broadband stations allowed me to better image these faults and examine the wider depth range of seismicity in this region of Oklahoma. Providing accurate depth values for earthquakes in regions of hypothesized induced seismicity has always been important, but gains further meaning in Oklahoma as differences between seismicity occurring in very shallow (<2 km depth) formations and within the crystalline basement could lead to different interpretations of the main driving stress change inducing said seismicity, namely the difference between hydraulic fracturing and fluid injection induced seismicity (Bao and Eaton, 2016; Holland, 2013; Skoumal et al., 2015, 2018; Brudzinski, 2019; Fasola et al., 2019; Savvaidis et al., 2020). It is in these cases that having access to dense networks such as the CWDE nodal network that could rapidly be deployed become even more important as they can provide far greater depth certainty in specific regions of interest where such uncertainties could lead to possible errors in the implementation of mitigation tactics.

4.10. Summary

Here I presented the creation of the CWDE_BB earthquake catalog created using the combined YW and XR broadband networks and the CWDE_NDL detection list created using the CWDE nodal network. The CWDE_BB catalog represents the most complete catalog of events detected and located by an analyst for the time period that both the CWDE broadband and nodal stations were deployed, along with event detections and locations created using an STA/LTA automated detection algorithm for the rest of the time period of the experiment. The earthquakes and associated focal mechanisms described within the CWDE_BB catalog image primarily E-W oriented strike-slip faults with earthquake depths ranging from 2-8 km placing these events

within the crystalline basement formation of northern Oklahoma. The CWDE_NDL detection list was created using the 363 nodal stations deployed as part of the CWDE project with the local similarity technique. While the local similarity technique lacked the ability to locate the events it detected, its created CWDE_NDL detection list can serve as the starting point for other automated techniques such as back-projection to create a rapid earthquake catalog for the large-N style nodal deployment. Possible future endeavors to utilize the CWDE nodal deployment to test automated detection and location techniques such as template matching or machine learning based approaches would greatly benefit from utilizing the CWDE_BB catalog as a comparison or training catalog.

4.11. Data and Resources

All seismic data utilized in this study was collected as part of the Community Wavefields Demonstration Experiment (CWDE) which focused on using a combined broadband and large-N style nodal station deployment to image the full wavefield and provide training on deployment of nodal arrays in northern Oklahoma, and the Seismicity near the Nemaha fault in northern Oklahoma (OK_GAP) project. The CWDE project was conducted as part of the Incorporated Research Institutions for Seismology (IRIS) Wavefields Demonstration Community Experiment using Program for the Array Seismic Studies of the Continental Lithosphere (PASSCAL) and rented instruments. The Oklahoma Wavefields dataset has been assigned the following doi: 10.7914/SN/YW_2016. These data have been archived at IRIS repository PH5 under network code YW and can be accessed at (http://service.iris.edu/ph5ws/dataselect/1/). The CWDE_BB catalog information is stored as an Antelope database whose software can be obtained at www.brtt.com. Injection volume information for saltwater disposal wells in Oklahoma can be obtained from the Oklahoma Corporation Commission's online public database at

https://oklahoma.gov/occ/divisions/oil-gas/oil-gas-data.html. The Oklahoma Geological Survey (OGS) earthquake catalog information can be obtained from their public online website https://www.ou.edu/ogs/research/earthquakes/catalogs. All websites were last accessed in August 2021.

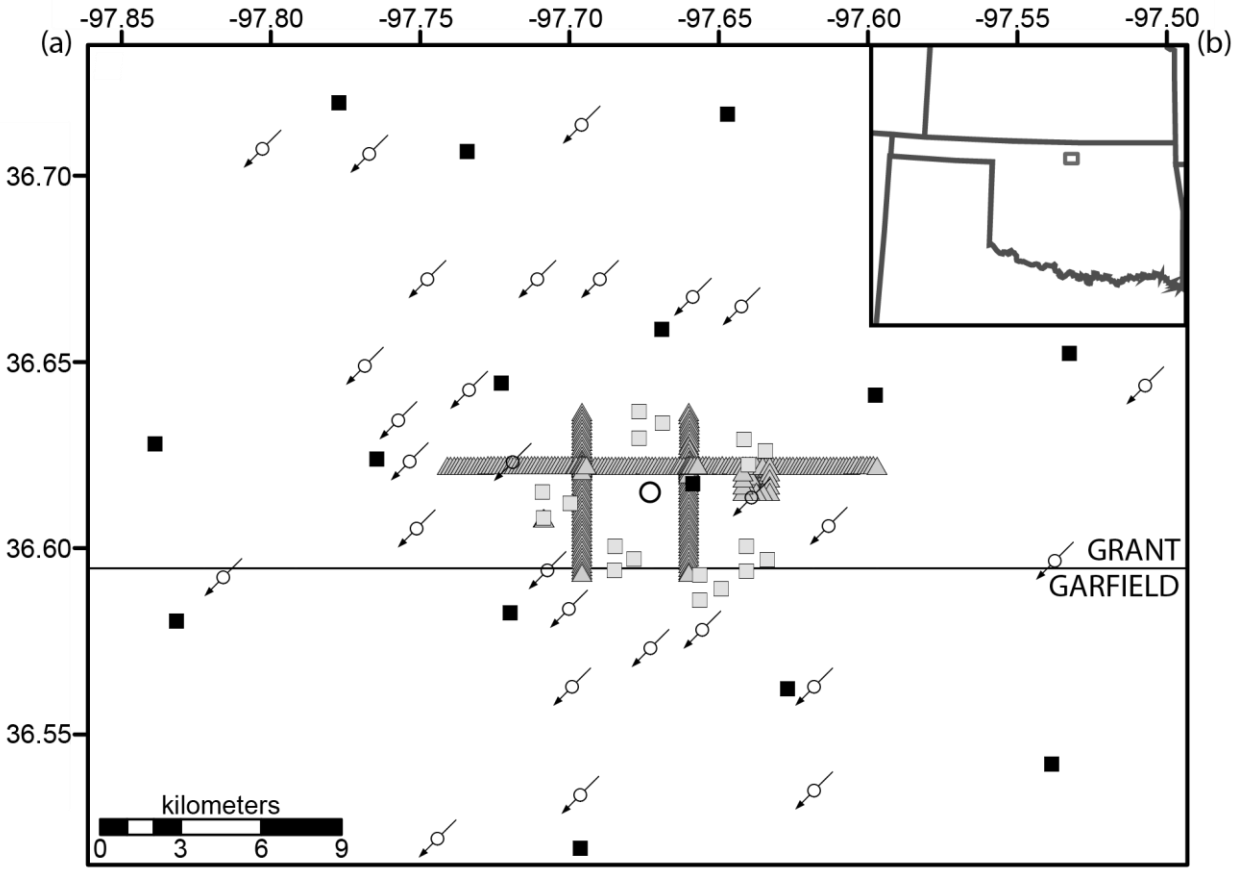

Figure 4.1. (a) Map view of the CWDE array (light gray) geometry in northern Oklahoma showing the locations of the nodal stations (triangles) and broadband stations (squares). Also shown are the locations of the XR broadband stations (black squares) that were used in this study and the locations of injection wells active during the period of study (arrows). The center point of the CWDE array is marked by the white circle. (b) Map view of Oklahoma showing the location of the CWDE array and region of study as a dark gray box.

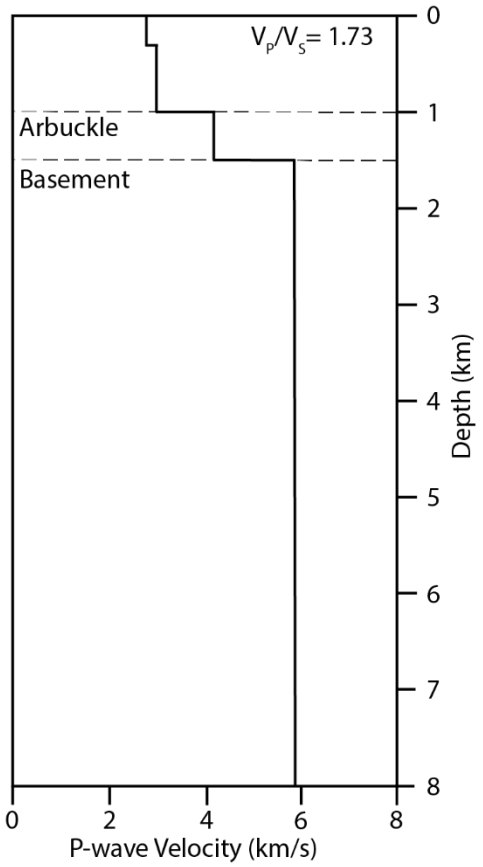

Figure 4.2. P-wave velocity model used in the creation of the CWDE_BB earthquake catalog showing the depth values of the Arbuckle and crystalline basement formations along with the used V_P/V_S ratio of 1.73 (Crain and Chang, 2018).

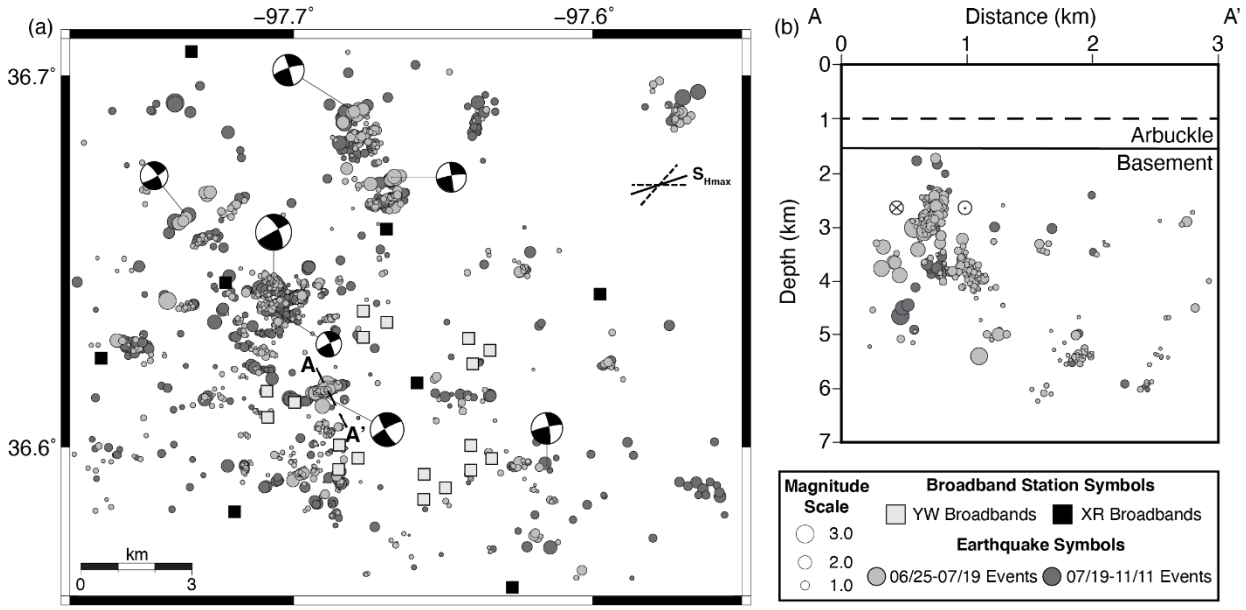

Figure 4.3. (a) Map view of the CWDE_BB earthquake catalog scaled by magnitude along with the locations of the YW and XR broadband stations used to create the catalog. Also shown are a selection of characteristic focal mechanism solutions for several of the imaged faults in the catalog. The regional maximum horizontal stress orientation (solid line labeled – Shmax) and optimal fault orientations (dashed lines) are shown in the top right. (b) Cross-sectional view of the CWDE_BB earthquake catalog with the fault motions of the main imaged fault shown.

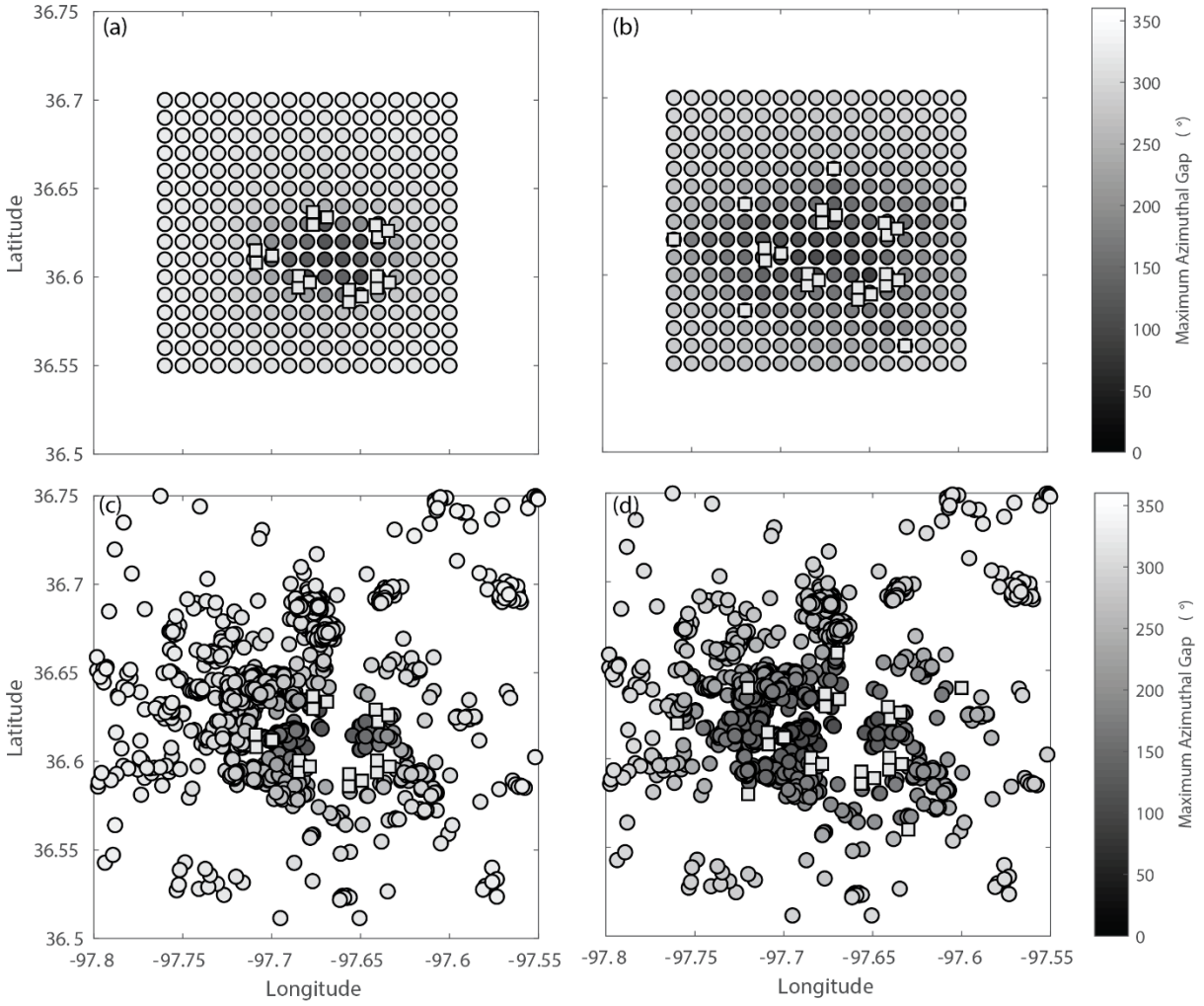

Figure 4.4. Map views of the CWDE and XR broadband stations (gray squares) showing the azimuthal gap values for theoretical and CWDE_BB catalog earthquakes (circles) using different network configurations. (a) Map of azimuthal gap values for a series of theoretical earthquake locations using only the CWDE broadband stations. (b) Map of azimuthal gap values for a series of theoretical earthquake locations using both the CWDE and XR broadband stations. (c) Map of azimuthal gap values for the CWDE_BB catalog earthquakes using only the CWDE broadband stations. (d) Map of azimuthal gap values for the CWDE_BB catalog earthquakes using both the CWDE and XR broadband stations.

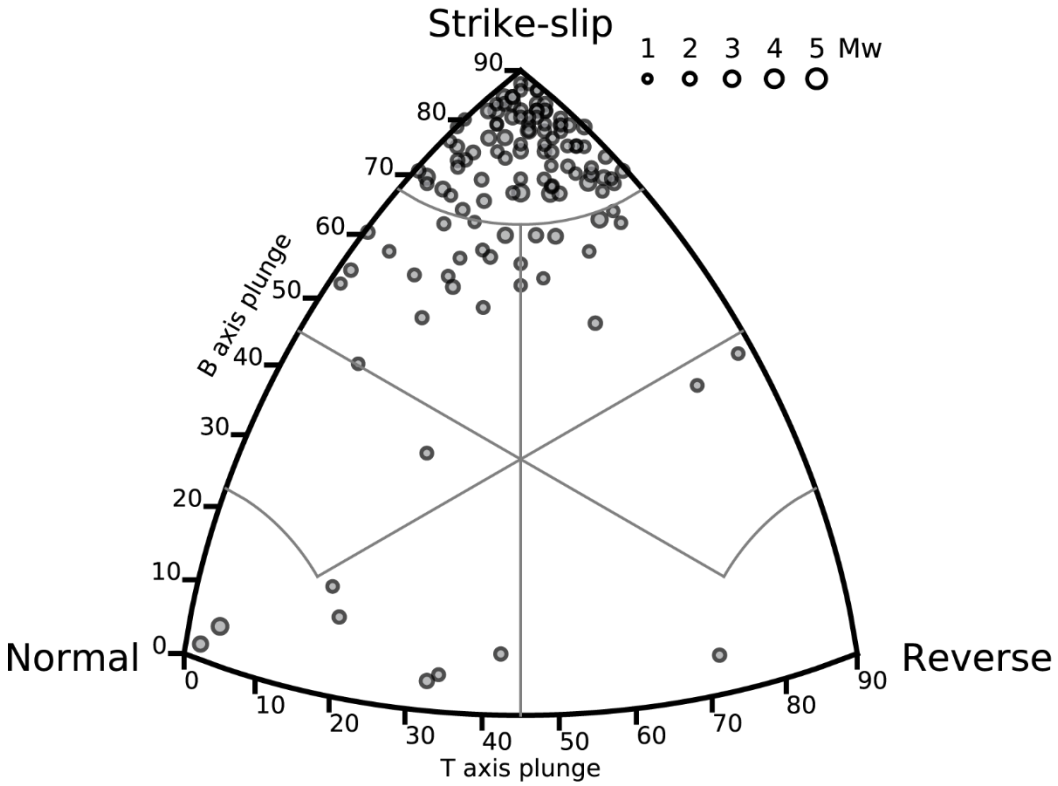

Figure 4.5. Ternary plot showing the distribution of focal mechanism solutions (circles) created from the M_L1.5+ earthquakes from the CWDE_BB catalog with their size scaled by magnitude.

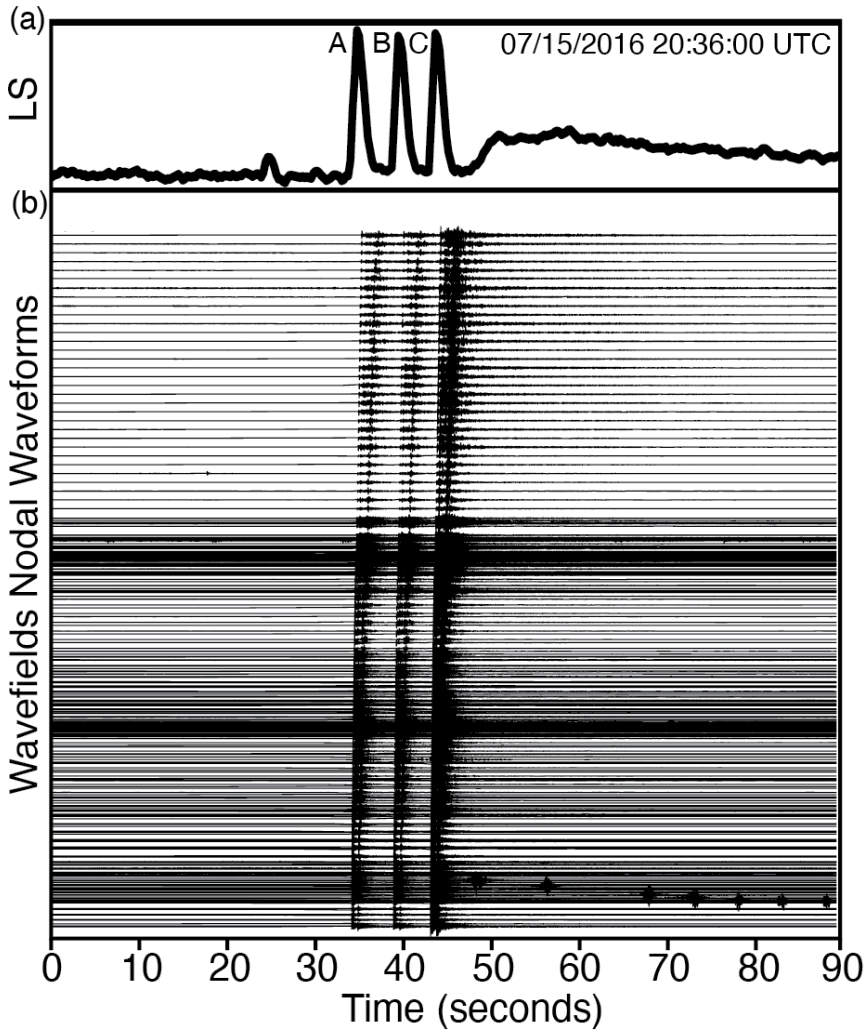

Figure 4.6. (a) View of the combined network local similarity value for the CWDE nodal array collected over a 90 second period showing three event detections marked as A, B, and C. (b) View of the raw waveform data collected for the 362 CWDE nodal stations active during this time period showing signal correlation across the network matching the marked event detections collected by the local similarity technique.

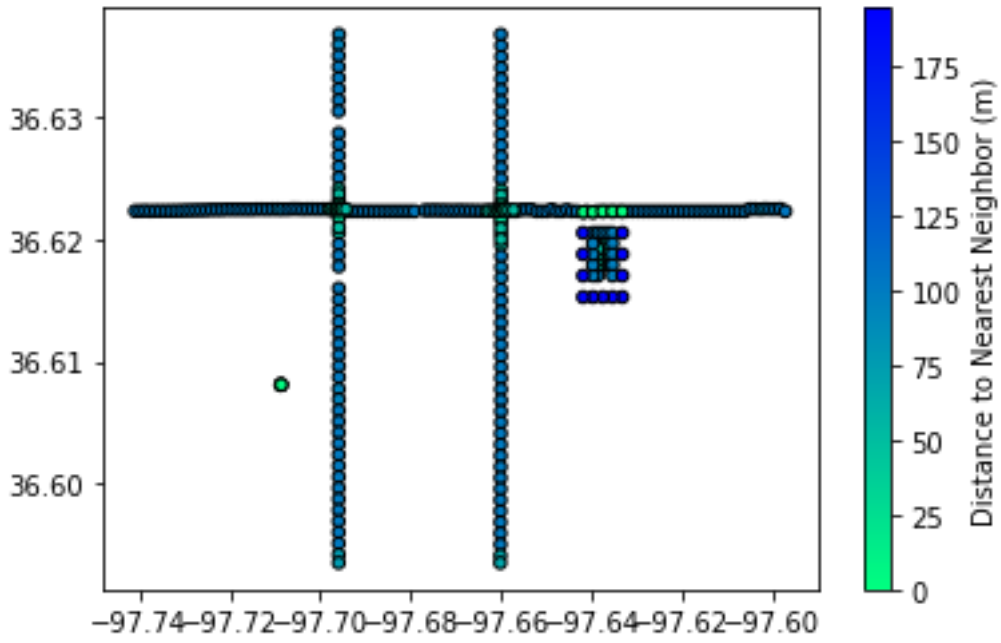

Figure 4.7. Map view of the CWDE nodal station network with each station colored by the distance from each individual station to its nearest neighbor station.

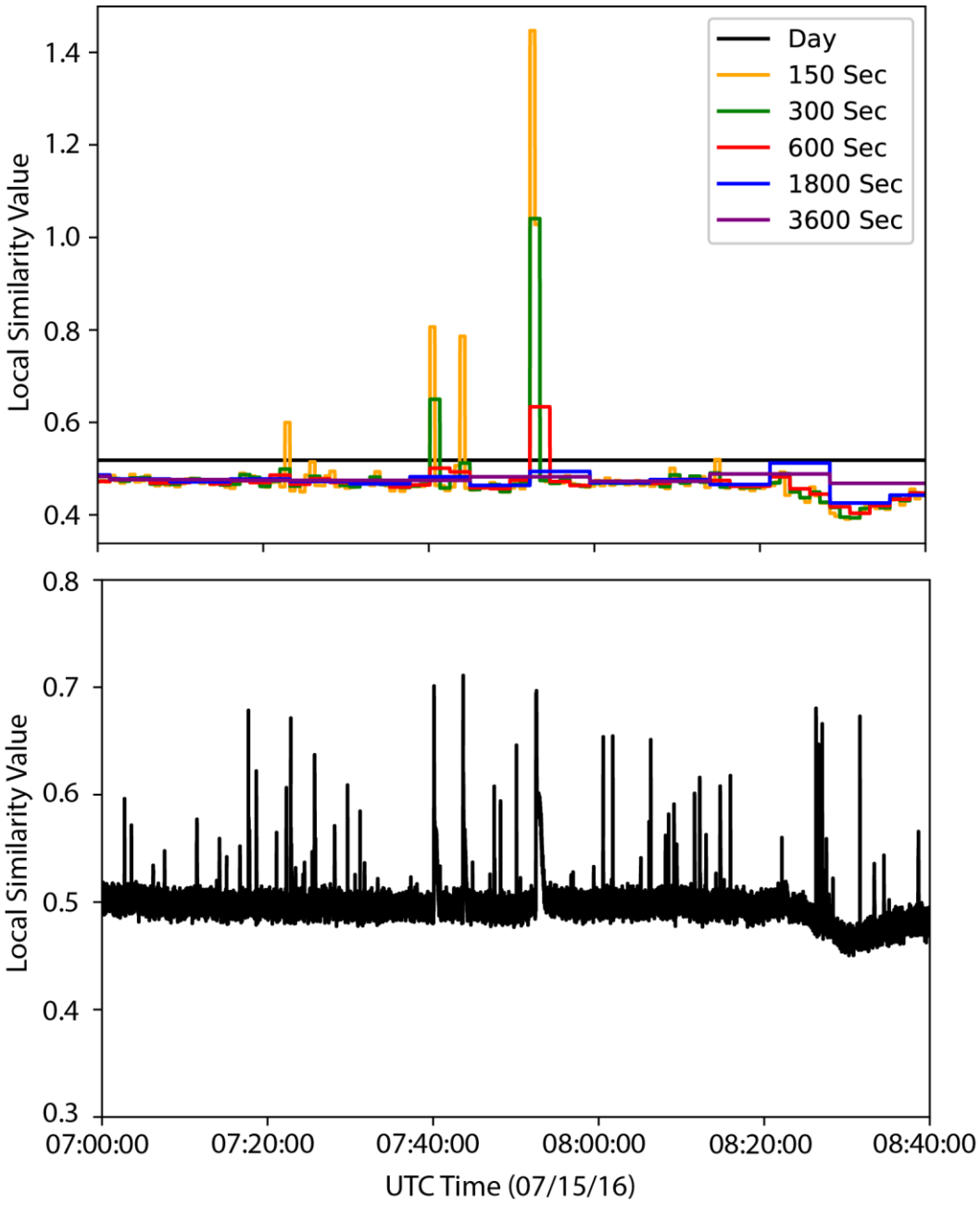

Figure 4.8. (top) Record of the "on" threshold values calculated over a 100-minute period using variable sliding time windows to calculate the 10*MAD (median absolute deviation) value ranging from 150 sec to an entire day. (bottom) Record of the raw local similarity values over the same 100-minute period.

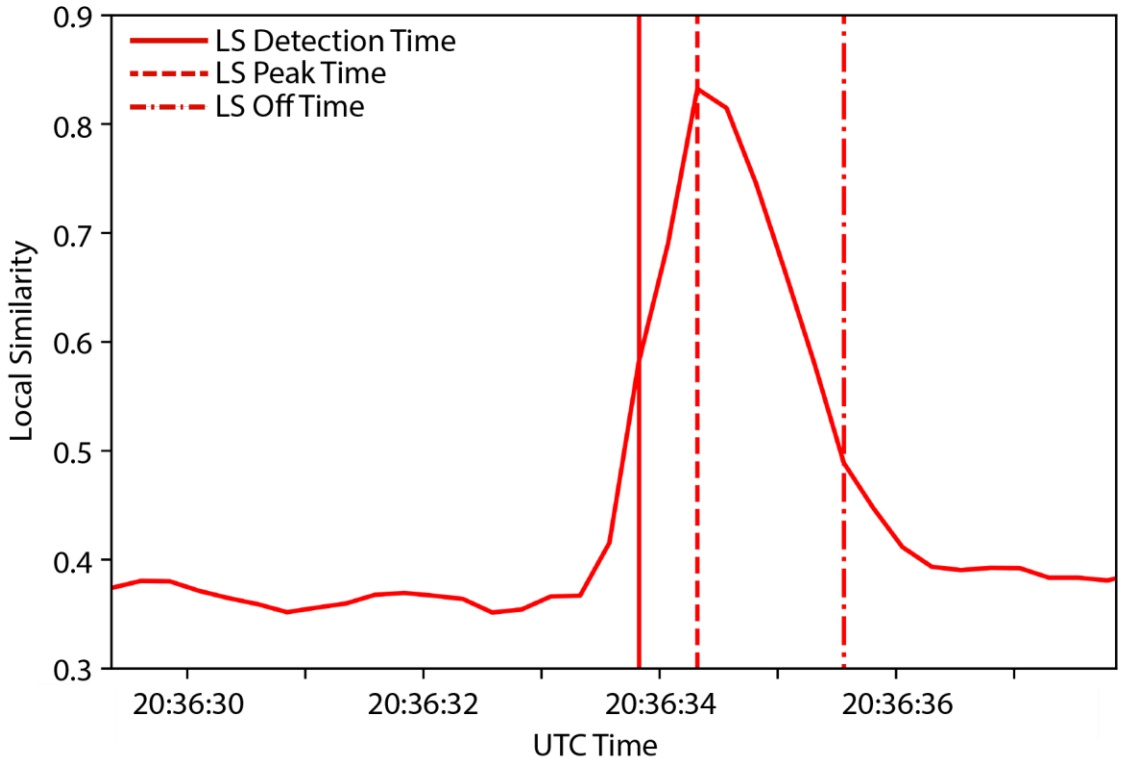

Figure 4.9. View of the local similarity results recorded by the DPZ channel for an event detected on 07/15/16 at approximately 20:36:34 UTC. The solid line represents the recorded LS detection time, the dashed line the LS peak time, and the dash-dot line the LS off time.

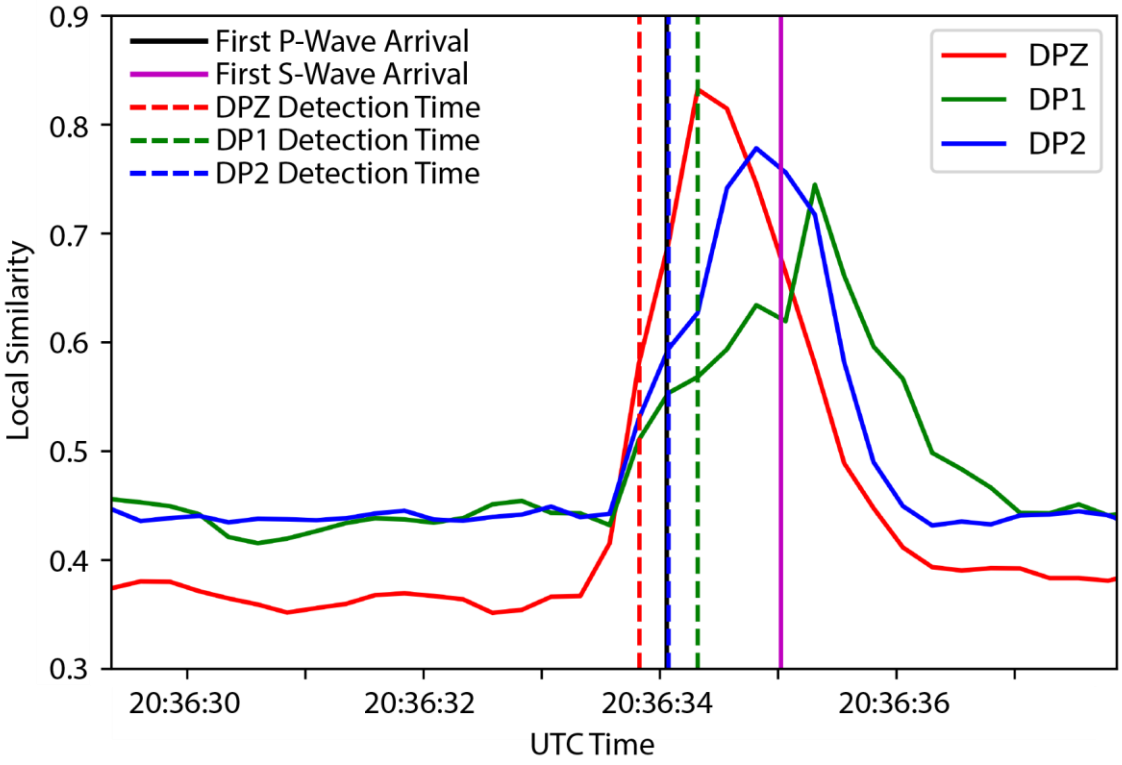

Figure 4.10. View of the local similarity results recorded on all three channels of the CWDE nodal station network for an event recorded on 07/15/16 at approximately 20:36:34 UTC. The colored dashed lines represent the LS detection times recorded for each individual channel. The first P-wave (black line) and S-wave (magenta line) arrival times recorded on the broadband stations are also shown.

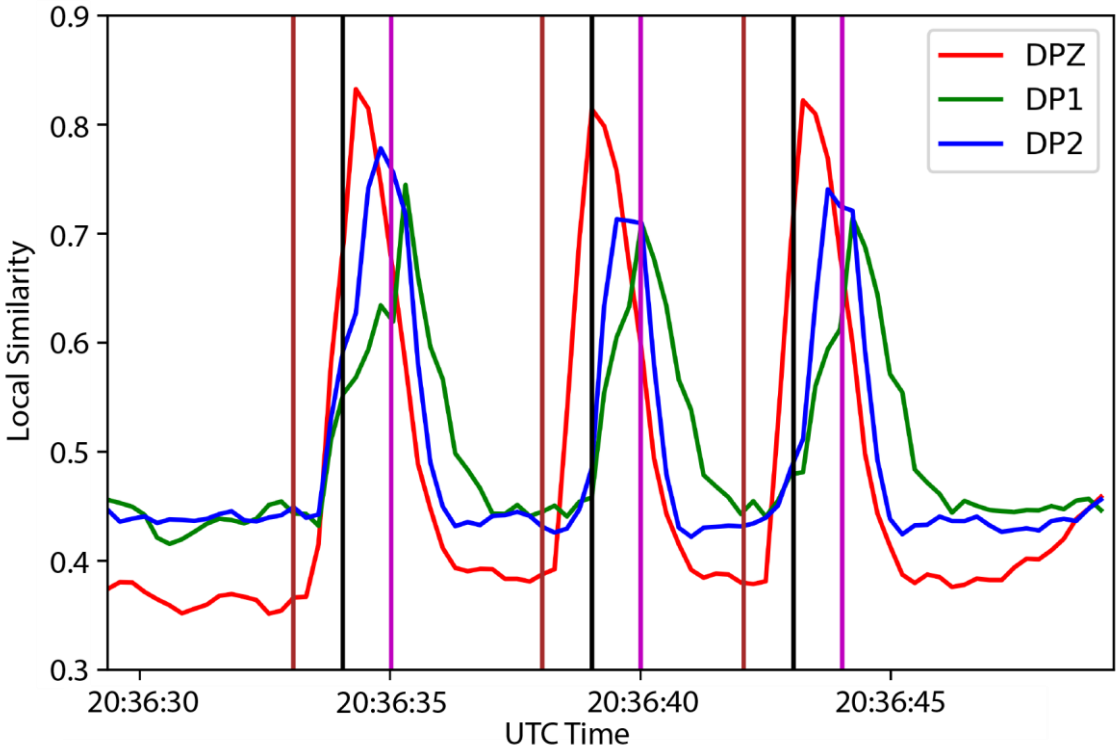

Figure 4.11. View of a set of the three events detected using local similarity which were shown in Figure 4.6 with the local similarity results from each of the individual channels shown. The arrival times of the first P-waves (black lines) and S-waves (magenta lines) to reach the CWDE broadband stations for each event are also shown. The origin times reported in the CWDE_BB catalog associated with each event detection (brown lines) are also shown.

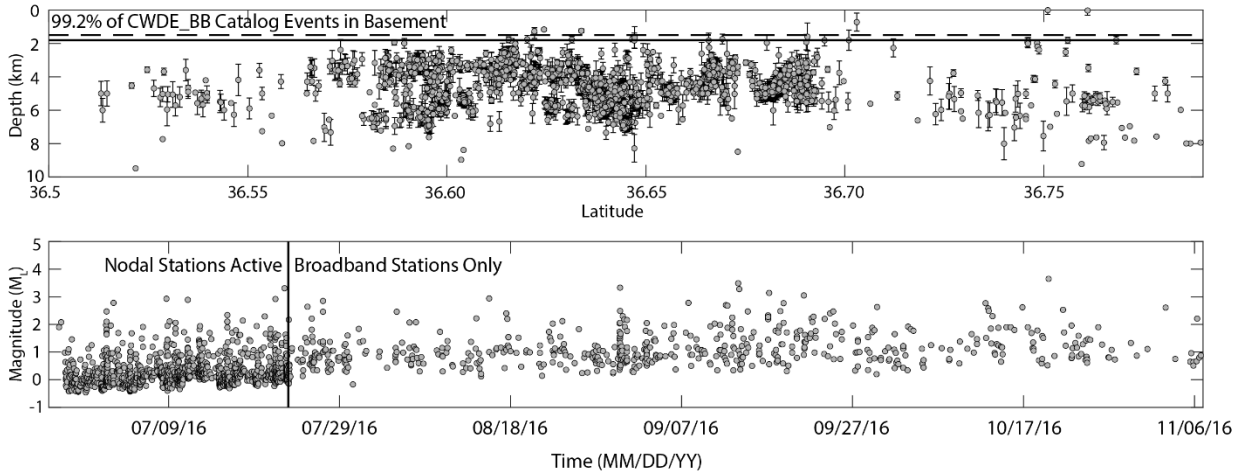

Figure 4.12. (top) N-S Cross-sectional view of the CWDE_BB catalog with the depth errors associated with each earthquake shown. Also shown are the depths (relative to sea level) of the Arbuckle (dashed line) and crystalline basement (solid line) formation tops. (bottom) Magnitude vs. time plot of the CWDE_BB catalog with the time delineation between the period when the nodal stations were active and when they were removed shown by the vertical black line.

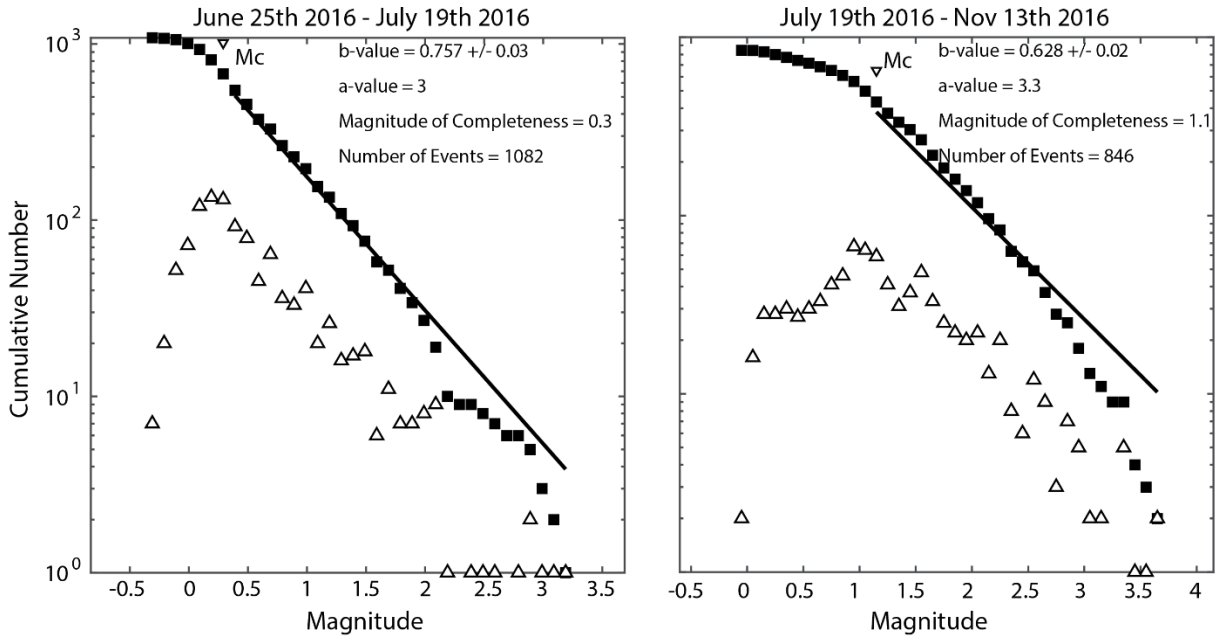

Figure 4.13. (left) Plot showing the calculated b-values and magnitude of completeness (Mc) values of the CWDE_BB catalog for the period when the nodal stations were active. The upwards pointing triangles represent the number of earthquakes in the CWDE_BB catalog associated with each magnitude bin (bin size=M_L0.1). The black squares represent the total number of events in the CWDE_BB catalog with a magnitude greater than or equal to the magnitude bin value. The downwards facing triangle represents the Mc value. The solid line represents the 90% goodness-of-fit approximation whose slope is equal to the b-value. (right) Plot showing the calculated b-values and magnitude of completeness values of the CWDE_BB catalog for the period when the nodal stations were inactive using the same symbology as the left panel.

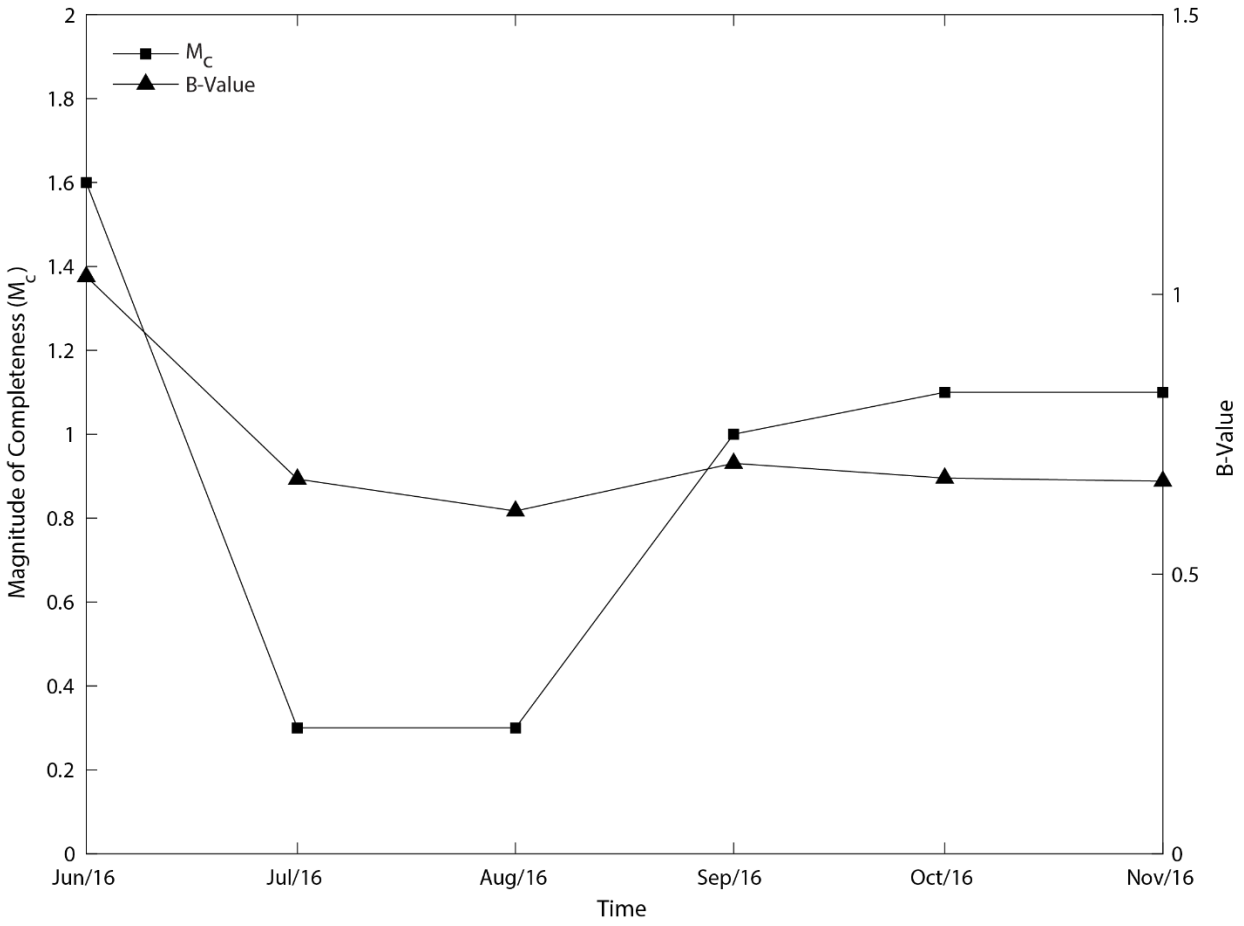

Figure 4.14. Plot showing the evolution of the overall CWDE_BB catalog magnitude of completeness (Mc) and b-values. For each temporal bin (bin size=one month) only the events occurring during or prior to that month were used in the calculation of the Mc and b-values for the catalog.

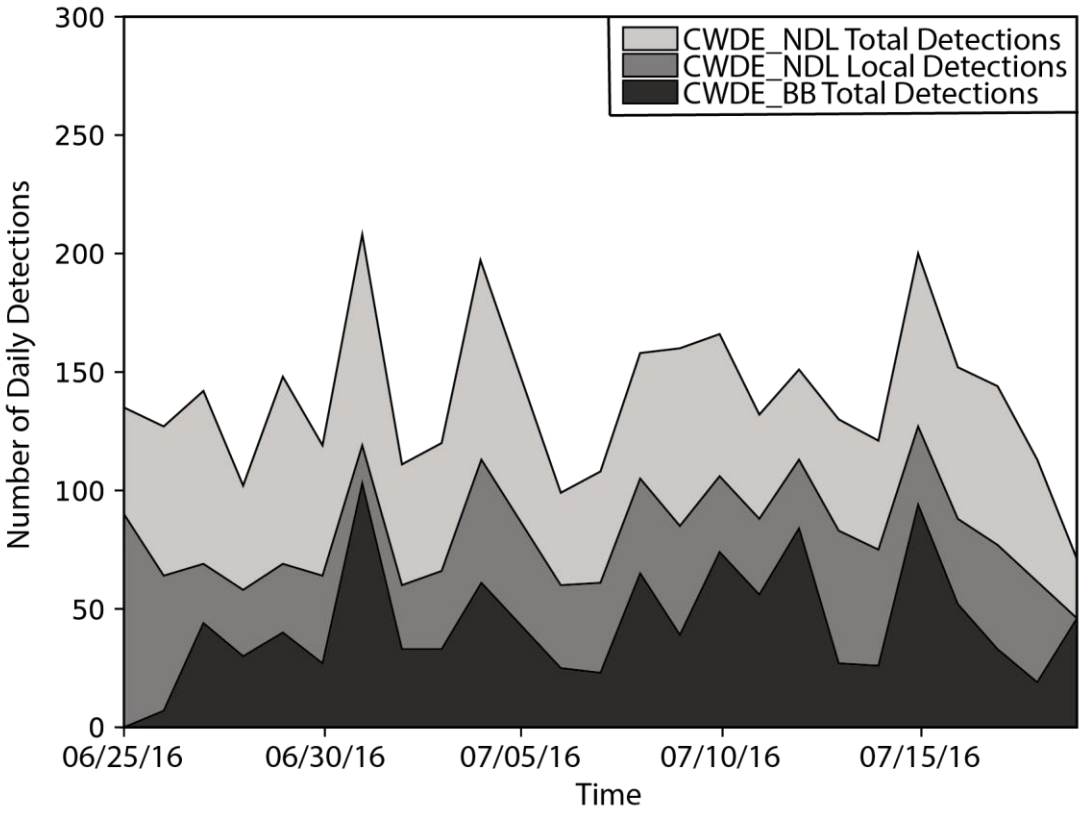

Figure 4.15. Temporal view of the number of detections collected every day over the course of the nodal deployment within the CWDE_BB catalog (dark gray) and CWDE_NDL detection list (light gray). Also shown are the number of detections recorded each day within the CWDE_NDL detection list whose S-minus-P time values match those expected of a local distance event from the center point of the CWDE array.

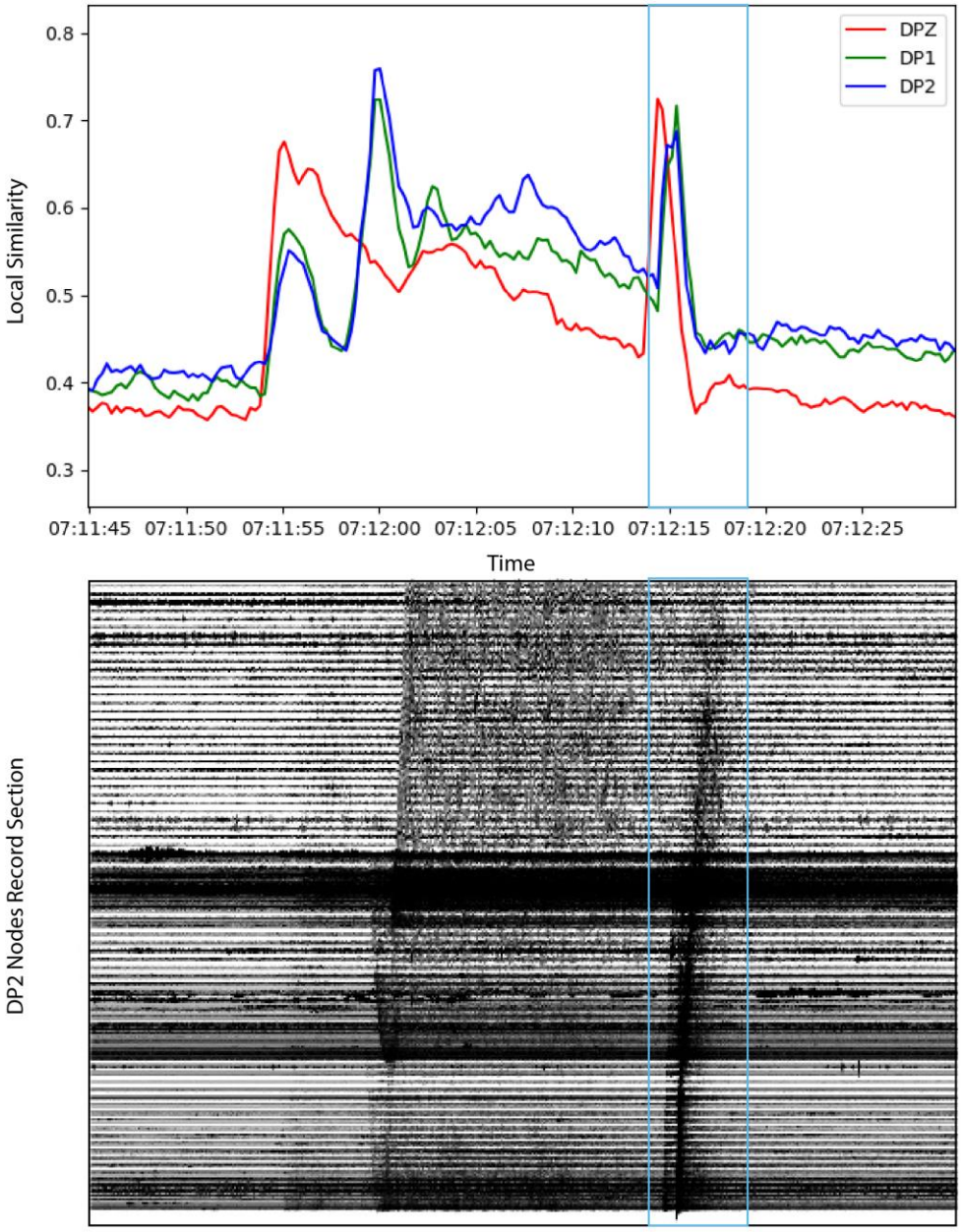

Figure 4.16. (top) Local similarity results collected on the three channels of the CWDE nodal network from a period where an event was recorded within the CWDE_BB catalog, but not within the CWDE_NDL detection list. (bottom) Seismic record collected from the CWDE nodal stations DP2 channel over the same period. The cyan box in both panels represents the arrival of the signal associated with the event recorded within the CWDE_BB catalog.

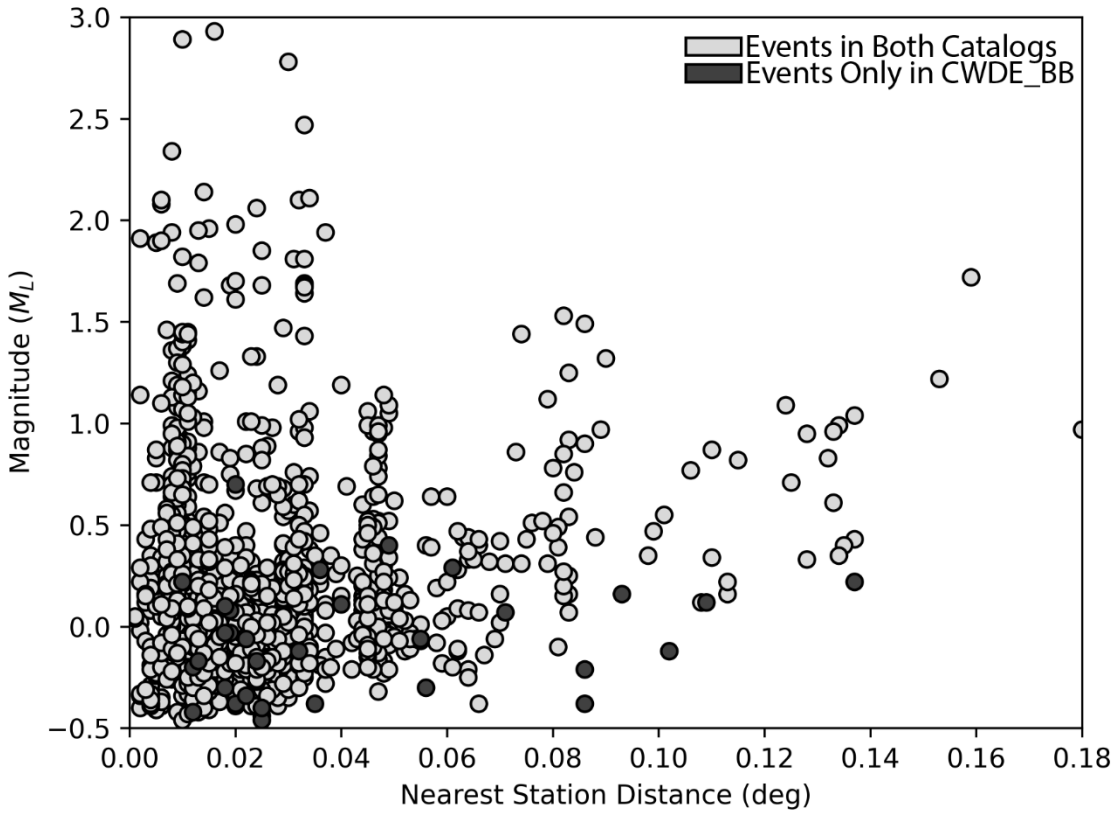

Figure 4.17. Plot showing the epicentral distance from each earthquake to the nearest recording CWDE broadband station vs. their recorded magnitude that is in both the CWDE_BB catalog and CWDE_NDL detection list (light gray circles) along with those events that are only in the CWDE_BB catalog (dark gray circles).

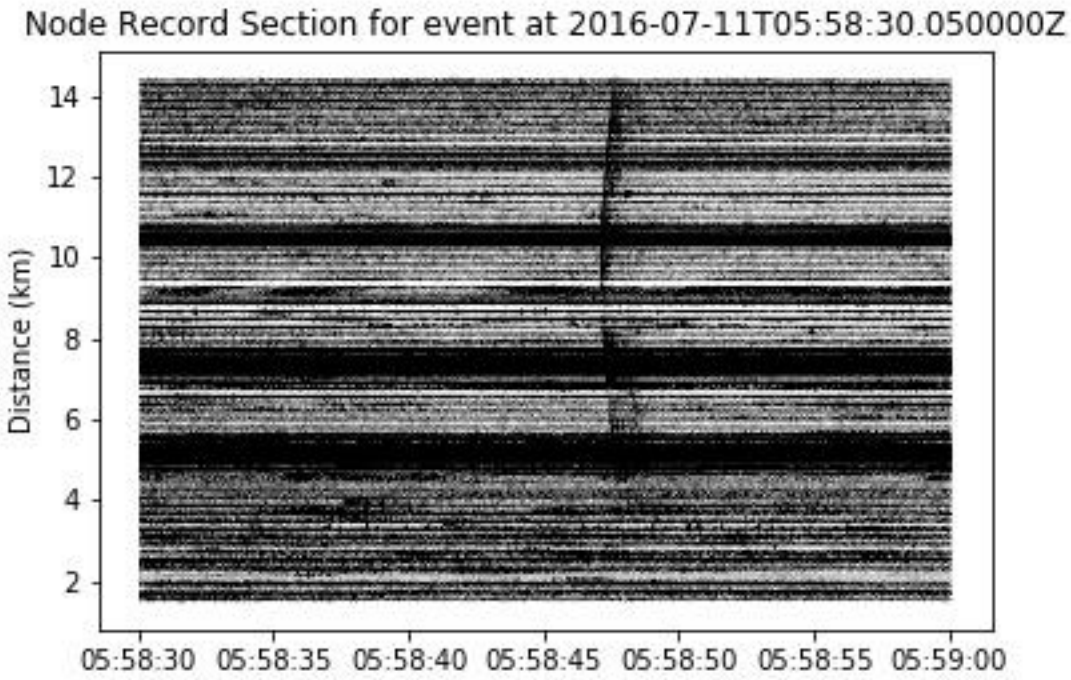

Figure 4.18. Seismic record section showing waveforms recorded on the DPZ channel of the CWDE nodal station network where local similarity detected an event which was not recorded within the CWDE_BB catalog.

5. DISCUSSION AND CONCLUSIONS

Here I present my work studying the Fort Worth Basin (FWB), located in north Texas, and the Community Wavefields Demonstration Experiment (CWDE) data set, located in northern Oklahoma, relating the seismicity occurring in both regions to fluid injection activities.

In the FWB study region my efforts focus on combining catalogs collected by researchers at Southern Methodist University (SMU) and relocating seismicity occurring within the basin covering the years 2008-2020. This work, also published as Quinones et al. (2019) provides the most comprehensive catalog of the induced seismicity occurring within the FWB, named the North Texas Earthquake Study (NTXES) catalog. The NTXES catalog is used in a basin-wide analysis of the seismicity to determine whether the seismicity has strong spatiotemporal links with the ongoing fluid injection activities also occurring within the basin; the study also expands the catalog such that the migration of seismicity through the basin over the full period of injection is clear, rather than solely focus on the five named sequences and faults that were targets of previous peer-reviewed publications. The increase in the completeness of the NXTES catalog as a whole and for each sub-group of the catalog is reflected in the lower magnitude of completeness values when compared to the regional ComCat calculated value. Further uses of the NTXES catalog involves the creation of a focal mechanism catalog which in combination with the earthquake locations themselves allows for interpretation of seismically active fault structures and the active local stress field. It is important to note that publication of all methods, 1D velocity models and attenuation relationships used for the NTXES locations and magnitudes

allows the NTXES catalog and all future events located by the SMU NTXES seismic network to enter the historic Texas earthquake catalog established by the Texas Seismological Network (TexNet); the work reflects a large collaborative effort extending over a decade.

The FWB seismicity contains multiple examples of faults reactivated far from fluid injection sources. The Dallas-Irving earthquake sequence represents the largest magnitude and highest risk example, and I study these events to answer lingering questions pertaining to the source of the stress change needed to have induced slip at far-field distances (>15 km). This work comprises creating a more high-resolution version of the Dallas-Irving using double-difference techniques to first better describe the source fault of the sequence and to examine the spatial extent of the seismicity. Another aspect of the seismic activity associated with the Dallas-Irving sequence is the actual start date of seismic activity associated with the Dallas-Irving source fault which within the NTXES and ComCat catalogs remains unclear. Due to budgetary and instrument inventory constraints, the monitoring and recording of seismicity has typically failed to capture the onset of seismicity along specific active faults in the FWB including the Dallas-Irving sequence. To rectify this issue, I set out to utilize the *EQcorrscan* template matching codes to better estimate the onset time of seismicity along the Dallas-Irving source fault. The template matching results presented in this dissertation and conducted by M. Brudzinski (in Hennings et al., 2021) show events associated with the Dallas-Irving fault may have begun as early as February 2014 or November 2013, respectively. However, the major focus of the work involving the Dallas-Irving sequence is the creation of a basin-wide 3D fully coupled geomechanical model which models the fluid injection associated pore fluid pressure and poroelastic stress changes within the FWB. The goal of this modeling effort is to better determine the spatiotemporal history of these injection associated stress changes and understand

how potentially far-field stress changes have affected the Dallas-Irving source fault and potentially triggered seismicity. The modeling results show that peak basin-wide injection associated stress changes occur within those portions of the basin where injection activities were highest, exceeding 1 MPa in the case of pore fluid pressure changes and 0.1 MPa in the case of poroelastic stress changes. The injection associated stress changes are far lower at the Dallas-Irving site, being lower than 0.01 MPa in the case of both stress changes, however the modeled stress changes arrived at the Dallas-Irving fault in 2008 far before the onset of seismicity.

The final portion of my work focuses not on the FWB data but rather the data collected in northern Oklahoma as part of the CWDE. My goal with utilizing this data set was to examine how a modern event detection technique could be implemented into the catalog creation workflow when studying a data set created using a modern large-N style array of stations. The specific technique I use is the local similarity technique, which is a waveform correlation-based approach to detecting possible event signals that has the advantage of not requiring *a priori* knowledge of the geology or physics of the study region. Using the local similarity technique, I create an event detection list which I then compared to an earthquake catalog I created using traditional analyst-based techniques on the same data set. Ultimately, I explore the benefits, such as the ability to rapidly process large amounts of data and create an event detection list, and limitations, such as the inability to locate the detected events without the use of other phase analysis techniques, that utilizing the local similarity technique provides and examine the possible avenues of analysis that could convert the created local similarity event detection list into a full earthquake catalog.

Through the creation of the NTXES catalog I sought to test my hypotheses about the relationship between the seismicity occurring within the FWB and fluid injection induced

seismicity. My first hypothesis involving the NTXES catalog was related to the spatiotemporal relationship between the fluid injection and seismic activities occurring within the basin. By relocating the earthquakes in the NTXES catalog using the updated 1D velocity models I can greatly lower the location uncertainties of the events with the largest improvements coming in the lowering of median depth uncertainty values (0.29 km in the NTXES catalog). Through the creation of the NTXES catalog I am able to examine the spatiotemporal history of seismic activity within the FWB and found that there did appear to be a direct correlation between the basin-wide fluid injection activities in the basin and the basin-wide activation of faults, affirming the relationship noted in Hornbach et al. (2016) with a more limited catalog. Although, while peak seismic activity (2014-2015) recorded within the NTXES catalog is time lagged when compared to peak injection activities in the basin (2011-2012), this may be a further effect of the lack of monitoring capabilities in the basin from 2010-2014.

My second hypothesis focuses on determining whether the seismicity in the FWB was driven by injection associated stress changes if the seismicity shared those previously mentioned key characteristics of fluid injection induced seismicity. I find that most active faults in the NTXES catalog events, except for the most northeastern sequences, are in close proximity to active injection wells and are composed of earthquakes occurring within the crystalline basement when examining the earthquake spatial and depth distributions relative to the locations of fluid injection wells and the top of the crystalline basement formation, respectively. Additionally, prior analysis of the regional stress field and source faults of many of the sequences also showed that these faults are critically stresses within the active regional stress field and thus required only small increases in stress to slip (Quinones et al., 2018; Hennings et al., 2019).

However, the question of if there is a pattern in the spatial and magnitude of stress changes at the sites of the active faults in the FWB remains which also leads to the question that if this pattern exists at what stress change threshold do the faults slip. Original estimates made in Quinones et al. (2018) used focal mechanism derived stress tensor and fault plane solutions to solve for possible ranges of stress changes needed to induce slip. Those stress change estimates were far higher when compared to other estimates made using fault slip probability (Hennings et al., 2019) or examining hydrogeologic modeling results (Gao et al., 2021; Hennings et al., 2021). Researchers at the Bureau of Economic Geology also sought to examine the stress change values at various sequence sites in the FWB at the time of recorded fault activation and found values ranging from 0.05-0.31 MPa across the basin. The reason for this much higher estimate of needed stress changes to induce slip is due to small changes in the fault plane orientations leading to large changes in the estimated pore pressure increase needed to induce slip within the same stress field. The reported stress changes in Quinones et al. (2018) were average values over the full range of higher quality focal mechanism solutions, which were greatly affected by inclusion of possible fault plane orientations which required these larger stress changes to slip. Here, I re-examined those stress change estimates using the same focal mechanism derived stress tensor solution, but instead of using the full range of focal mechanism derived fault planes used only the fault plane interpretation created using the combined earthquake location and seismic reflection data. By using only that singular fault interpretation I lose the ability to report any uncertainty values associated with this value as all uncertainty is inherent with the interpretation of the fault plane. However, the estimated stress changes needed to induce slip (0.02 MPa and 0.26 MPa for the Venus and Dallas-Irving faults, respectively) are much lower in magnitude and

better match the modeled stress changes observed at both sequence sites within both my geomechanical model and the Gao et al. (2021) hydrogeologic model.

The final hypothesis related to the NTXES data set is related to those mentioned northeastern sequences in the FWB which lacked one key characteristic of fluid injection induced seismicity, that being the presence of nearby fluid injection well to the active fault. The hypothesis was that if injection activities continued over a long enough time span, then far-field injection associated stress changes should be expected and observable. To test this hypothesis, I focus my research efforts on the Dallas-Irving sequence and examining what possible stress changes could have induced slip on its source fault. The creation of the double-difference relocated Dallas-Irving catalog helps correct the earthquake locations of the earliest events associated with the sequence. These earliest earthquake locations have much higher location uncertainties and tend not to fall along the imaged fault plane the other earthquakes create due to these earliest events being located using only regional station data because the local network was not yet installed. Ultimately, most of these early earthquakes are relocated onto the main cluster of earthquake locations associated with the Dallas-Irving fault and uniquely images this NE-SW trending steeply dipping normal fault.

The goal of creating the 3D geomechanical model of the FWB is to try and answer my previously stated hypothesis relating to the far-field injection associated stress changes and their possible effects on faults which do not lie near fluid injection wells. Because there are no well logs in the area around the Dallas-Irving sequence area to pull stress state data from, I must rely upon geomechanical modeling results to try and estimate the level of stress increases observed at the Dallas-Irving site. The modeling results describe the spatiotemporal evolution of the basin-wide stress changes where peak pore fluid pressure and poroelastic stress change are highest in

those portions of the basin where injection activities were most concentrated. However, over the time span of the model (2005-2019) I observe that both the pore fluid pressure and poroelastic stress changes associated with injection of fluids into the Ellenburger formation spread all throughout the basin with the increases in both stress changes leading to instability in the basement which makes fault slipping easier. Thus, while the modeling results suggest that the injection associated stress changes at the Dallas-Irving sequence site are small, they are still observable. The two models which best serve as comparisons to the results presented here are the models created by Zhai and Shirzaei (2018) and Gao et al. (2021). The major differences between the three models are that, while the Zhai and Shirzaei model was also fully coupled, it did not incorporate the full fault map and geologic parameters now available, and while the Gao et al. model did incorporate those same input parameters, it did not solve for poroelastic stress changes. Regardless both models showed significantly lower injection associated stress changes at the Dallas-Irving sequence site relative to the portions of the basin where injection activities were highest. Both models also record higher stress changes at the Dallas-Irving site when compared to my modeling results, which reflects the uncertainty on absolute value imparted by differences in input parameters and physics implementation within the different software. Ultimately, in all modeling cases seismicity associated with the Dallas-Irving sequence only began following the arrival of the injection associated stress changes reaching the fault. I conclude that those stress changes imparted by far-field waste fluid injection were sufficient and the likely causal mechanism of Dallas-Irving seismicity.

To fully address the relative timing between fluid injection induced stress changes on the Dallas-Irving fault, the true onset time of fault activation was required. I use the *EQcorrscan* template matching codes and regional station data from the year prior to the installation of the

local monitoring network around the Dallas-Irving sequence. In the process I find that seismicity associated with the Dallas-Irving sequence may have begun as early as February 2014, almost a full ten months prior to the earliest events included in the NTXES catalog for the Dallas-Irving sequence. This result is different from previous work. H. DeShon and C. Hayward, using S-P times and the review of the ComCat, had previously identified an April 2014 felt earthquake as likely on the Dallas-Irving fault (unpublished, personal communication). Template matching efforts conducted by M. Brudzinski using distant station data (>150 km) found events in 2013 (Hennings et al., 2021). However, there is some uncertainty in the template matching detection lists associated with the earliest events as those events could easily be associated with the DFW airport fault, which was also seismically active during the same period (Ogwari et al., 2018). Nevertheless, even if the earliest events from both template matching catalogs were indeed Dallas-Irving events, the modeling results suggest that the injection associated stress changes had already reached the fault several years prior in 2008. Therefore, again, seismicity began only after the fault had begun to be additionally stressed due to regional fluid injection activities.

The final portion of my work described in this dissertation focuses on the CWDE data set and my efforts to utilize the local similarity technique to create a starting event detection list rapidly and efficiently for a large-N style array data set. The work in this chapter revolves around the question about how significantly do modern event detection techniques aid in the analysis of modern network deployments. The CWDE array provided an excellent study site due to its inclusion of both a large-N style nodal array of over 360 stations and a more traditional broadband array deployed within a small spatial footprint over the same period (Sweet et al., 2018). To study this data set I first create a benchmark catalog using more traditional analyst-based techniques utilizing the broadband station data which also serves as the first recorded

catalog of seismicity associated with the CWDE data set. However, most of my work focuses on using the local similarity event detection technique on the nodal array to assess how well the technique performs in comparison to the benchmark catalog. The nodal array data set is unfeasible for an analyst to have to work through manually and so to realistically be able to analyze the data set in a reasonable amount of time an automated approach must be utilized. It's here that the local similarity technique performed very well with the full processing of all three channels of the 363-station nodal array being completed in just under 48 hours. The results of the local similarity analysis are a set of singular network wide valuations of signal coherency across the entirety of the nodal array which could be used to detect the arrival of earthquake signals. Analysis of the local similarity results with regards to the actual waveform data reveals that the local similarity technique could be used to determine arrival times of the first P- and S-waves to reach the array. Thus, while the local similarity technique has no built-in methodology for the estimation of an earthquake location, I could use the P- and S-wave arrival times to estimate an event-to-array distance value using S-minus-P times. The full event detection list created using the local similarity technique is separated into those events which were considered local (<15 km event-to-array distance) and regional, and the local event detection list is compared to the benchmark catalog. The comparison of both techniques revealed that the local similarity technique successfully detected 97.0% of all events in the benchmark catalog, and re-assessment of the missed detections revealed that they were due to needed strict definitions of event detection parameters rather than a total inability of the local similarity technique to identify those events.

5.1. Conclusions

The NTXES catalog as the most complete record of seismicity occurring in the FWB represents a great starting point for possible future analysis of long-spanning induced seismicity. The NTXES catalog is created using traditional analyst-based approaches to earthquake catalog creation and thus possible future avenues involve the incorporation of more modern event detection and phase picking approaches to expand upon the created catalog. Template matching efforts to help determine the onset times of seismicity along the active faults in the FWB and to help in the detection of missed small magnitude events are a likely next step for analyzing seismicity in the basin. Additionally, machine learning based phase picking approaches are also being explored using the FWB data and are likely to yield higher event counts associated with the major active faults in the basin. However, in both cases the incorporation of a proper event association and earthquake location step will need to be incorporated into the workflow for these event detection techniques to yield full earthquake catalogs which could be compared to the compiled NTXES catalog. As of August 2021, the Azle fault has not generated earthquakes in a few years, the Dallas-Irving fault has not hosted an event since 08/09/2020 and the Venus area has had 12 events in 2021. The continued monitoring of seismicity in the FWB planned for 2021-2023 through TexNet should yield new insight into both the lifespan of induced seismicity sequences, and provide insight into how and why induced seismic sequences may “die”. In fact, the FWB seismicity could inform us about the effectiveness of injection mitigation techniques as due to a variety of factors injection activities have gradually been decreasing since their peak in 2012 and likewise seismicity rates have continuously decreased following their peak in 2015.

For monitoring, the created event detection list of the CWDE data set serves as a great starting point, but further analysis steps need to be undertaken to convert these results into a full

earthquake catalog. Thus, future work associated with the generated local similarity event detection list should focus on using the found phase arrival times as targeted time windows for a phase picking algorithm to pick arrival times for each individual station rather than for the entire network. Then, work could be done to either use traditional earthquake location algorithms or even algorithms such as back-projection to gather actual estimates of earthquake locations for these events. As for magnitude estimates, it would be simplest to use an amplitude-based magnitude calculation approach following the determination of individual phase arrival times for each nodal station. Further analysis of the CWDE data set may be extended to other event detection and phase picking techniques and algorithms which could either seek to take advantage of the benchmark catalog for approaches such as template matching or machine learning or instead focus entirely on the nodal array data and benchmark their results on the local similarity results instead. Ultimately, the CWDE data set remains an excellent data set for the purposes of testing a variety of event detection and phase picking techniques which could be incorporated into the workflow analysis of any future network deployment involving the installation of large-N style broadband and nodal station arrays.

APPENDIX

Appendix information is listed here and the associated files for each appendix are available in electronic formats upon request.

Name: Louis Quinones
Email: louquinones@outlook.com

Appendix A.1. NTXES catalog information including the origin time of each event, the event location, location uncertainty values (reported as 68% confidence error ellipses) and named velocity model used for each event location (2008-2018).

Appendix A.2. Table views of the four velocity models used to locate the events in the NTXES catalog showing both the P- and S-wave velocity values.

Appendix B.1. Catalog of the double-difference relocated Dallas-Irving earthquakes in the format described in the *hypoDD* and *tomoDD* manuals.

Appendix C.1. Earthquake catalog created for the Community Wavefields Demonstration Experiment (CWDE) data set using the broadband instruments containing information about the origin time of each event, the event location, and uncertainty values (reported as 68% confidence error ellipses).

Appendix C.2. List of detected events for the Community Wavefields Demonstration Experiment (CWDE) using the local similarity method with the associated P- and S-wave arrival times for each event listed and whether the event was a local or regional event listed.

Appendix C.3. Locations, names, and monthly injection volumes from fluid injection wells active in Grant and Garfield counties during the period the CWDE array was active.

Appendix C.4. Catalog of the focal mechanisms created using the CWDE_BB earthquake catalog.

BIBLIOGRAPHY

Alt, R. C., & Zoback, M. D. (2017). In situ stress and active faulting in Oklahoma. *Bull. Seism. Soc. Am.*, *107*(1), 216-228.

Álvarez-Gómez, J. A. (2019). FMC—Earthquake focal mechanisms data management, cluster and classification. *SoftwareX*, *9*, 299-307.

Atkinson, G., & D. Boore (2013). The attenuation of Fourier amplitudes for rock sites in eastern North America, *Bulletin of the Seismological Society of America*, *104*, 513-528. https://doi.org/10.1785/0120130136.

Bao, X., & Eaton, D. W. (2016). Fault activation by hydraulic fracturing in western Canada. *Science*, *354*(6318), 1406-1409.

Bardwell, G. E. (1966). Some statistical features of the relationship between Rocky Mountain Arsenal waste disposal and frequency of earthquakes. *The mountain geologist*.

Bear, J. (2013). Dynamics of fluids in porous media. *Courier Corporation*.

Beaucé, E., Frank, W. B., & Romanenko, A. (2018). Fast matched filter (FMF): An efficient seismic matched-filter search for both CPU and GPU architectures. *Seismological Research Letters*, *89(1)*, 165-172.

Bender, B. (1983). Maximum likelihood estimation of b values for magnitude grouped data. *Bulletin of the Seismological Society of America*, *73(3)*, 831-851.

Bolarinwa, O. J., & Langston, C. A. (2021). Calibrating the 2016 IRIS Wavefields Experiment Nodal Sensors for Amplitude Statics and Orientation Errors. *Bull. Seism. Soc. Am.*.

Bowker, K. A. (2007). Barnett shale gas production, Fort Worth Basin: Issues and discussion. *AAPG bulletin*, *91(4)*, 523-533.

Brudzinski, M. R., & Kozłowska, M. (2019). Seismicity induced by hydraulic fracturing and wastewater disposal in the Appalachian Basin, USA: A review. *Acta Geophysica*, *67*(1), 351-364.

Cao, A., & Gao, S. S. (2002). Temporal variation of seismic b-values beneath northeastern Japan island arc. *Geophysical research letters*, *29(9)*, 48-1.

Cardott, B. J. (2017). Oklahoma shale resource plays. *Oklahoma Geology Notes, 76(2)*, 21-30.

Chamberlain, C. J., Hopp, C. J., Boese, C. M., Warren-Smith, E., Chambers, D., Chu, S. X., ... & Townend, J. (2018). EQcorrscan: Repeating and near-repeating earthquake detection and analysis in Python. *Seismological Research Letters, 89(1)*, 173-181.

Chamberlain, C. J., Townend, J., & Gerstenberger, M. C. (2020). RT-EQcorrscan: Near-Real-Time Matched-Filtering for Rapid Development of Dense Earthquake Catalogs. *Seismological Society of America, 91(6)*, 3574-3584.

Chang, K. W., & Segall, P. (2016). Injection-induced seismicity on basement faults includ-ing poroelastic stressing. *Journal of Geophysical Research: Solid Earth, 121(4)*, 2708-2726. Doi:10.1002/2015JB012561.

Chang, K. W., & Yoon, H. (2020). Hydromechanical controls on the spatiotemporal patterns of injection-induced seismicity in different fault architecture: Implication for 2013–2014 Azle earthquakes. *Journal of Geophysical Research: Solid Earth, 125*, e2020JB020402. https://doi.org/10.1029/2020JB020402.

Chen, C., & Holland, A. A. (2016). PhasePApy: A robust pure Python package for automatic identification of seismic phases. *Seismol. Res. Lett., 87*(6), 1384-1396.

Cochran, E. S., Wickham-Piotrowski, A., Kemna, K. B., Harrington, R. M., Dougherty, S. L., & Peña Castro, A. F. (2020). Minimal clustering of injection-induced earthquakes observed with a large-N seismic array. *Bull. Seism. Soc. Am., 110*(5), 2005-2017.

Crain, K. D., & Chang, J. C. (2018a). Elevation map of the top of the crystalline basement in Oklahoma and surrounding states. *Open-File Report 1-2018*. www.ou.edu/ogs/publications/openfilereports.

Crain, K. D., & Chang, J. C. (2018b). Elevation and thickness of the Ordovician arbuckle group in Oklahoma and surrounding states. *Open-File Report 2-2018.* www.ou.edu/ogs/publications/openfilereports.

Davis, S. D., & Frohlich, C. (1993). Did (or will) fluid injection cause earthquakes?-criteria for a rational assessment. *Seismological Research Letters, 64(3-4)*, 207-224.

Deng, K., Liu, Y., & Harrington, R. M. (2016). Poroelastic stress triggering of the December 2013 Crooked Lake, Alberta, induced seismicity sequence. *Geophysical Research Letters, 43(16)*, 8482-8491.

DeShon, H. R., Hayward, C. T., Ogwari, P. O., Quinones, L., Sufri, O., Stump, B., & Magnani, M. (2018). Summary of the North Texas Earthquake Study Seismic Networks, 2013-2018. *Seismological Research Letters, 90*(1), 387-394. doi:10.1785/020180269.

Ellsworth, W. L. (2013). Injection-induced earthquakes. *Science, 341*(6142), 1225942. Doi:10.1126/science.1225942.

Ewing, T. (1990). Tectonic map of Texas, published by *Bureau of Economic Geology,* University of Texas, Austin, Texas.

Fan, Z., Eichhubl, P., & Newell, P. (2019). Basement fault reactivation by fluid injection into sedimentary reservoirs: Poroelastic effects. *Journal of Geophysical Research: Solid Earth, 124(7)*, 7354-7369.

Fan, W., & McGuire, J. J. (2018). Investigating microearthquake finite source attributes with IRIS Community Wavefield Demonstration Experiment in Oklahoma. *Geophys. J. Int.*, *214*(2), 1072-1087.

Fasola, S. L., Brudzinski, M. R., Skoumal, R. J., Langenkamp, T., Currie, B. S., & Smart, K. J. (2019). Hydraulic fracture injection strategy influences the probability of earthquakes in the Eagle Ford shale play of South Texas. *Geophys. Res. Lett.*, *46*(22), 12958-12967.

Frohlich, C. (2012). Two-year survey comparing earthquake activity and injection-well locations in the Barnett Shale, Texas. *Proceedings of the National Academy of Sciences*, *109*(35), 13934-13938. Doi:10.1073/pnas.1207728109.

Frohlich, C., Potter, E., Hayward, C., & Stump, B. (2010). Dallas-Fort Worth earthquakes coincident with activity associated with natural gas production, *Leading Edge, 29*, 270-275. doi:10.1785/0120100131.

Frohlich, C., Hayward, C., Stump, B., & Potter, E. (2011). The Dallas-Fort Worth earthquake sequence: October 2008 through May 2009, *Bulletin of the Seismological Society of America, 101*(1), 327-340. doi: 10.1785/0120100131.

Frohlich, C., DeShon, H., Stump, B., Hayward, C., Hornbach, M., & Walter, J. I. (2016). A historical review of induced earthquakes in Texas, *Seismological Research Letters, 87*(4), 1022-1038. doi:10.1785/0220160016.

Gao, R. S., Nicot, J. P., Hennings, P. H., La Pointe, P., Smye, K. M., Horne, E. A., & Dommisse, R. (2021). Low pressure build-up with large disposal volumes of oilfield water: A flow model of the Ellenburger Group, Fort Worth Basin, northcentral Texas. *AAPG Bulletin, (20,210,707)*.

Gibbons, S. J., & Ringdal, F. (2006). The detection of low magnitude seismic events using array-based waveform correlation. *Geophysical Journal International, 165(1)*, 149-166.

Geotechnical Corporation (1964). Deep-hole site report, Trigg no. 1, Dallas County, Texas, *Geotechnical Corporation Technical Report 64-100*.

Goebel, T. H. W., Weingarten, M., Chen, X., Haffener, J., & Brodsky, E. E. (2017). The 2016 Mw5. 1 Fairview, Oklahoma earthquakes: Evidence for long-range poroelastic triggering at> 40

km from fluid disposal wells. *Earth and Planetary Science Letters, 472,* 50-61. Doi:10.1016/j.epsl.2017.05.011.

Goebel, T. H., & Brodsky, E. E. (2018). The spatial footprint of injection wells in a global compilation of induced earthquake sequences. *Science, 361(6405),* 899-904.

Gono, V., Olson, J. E., & Gale, J. F. (2015). Understanding the correlation between induced seismicity and wastewater injection in the Fort Worth Basin. In *49th US Rock Mechanics/Geomechanics Symposium*. American Rock Mechanics Association.

Grigoli, F., Cesca, S., Vassallo, M., & Dahm, T. (2013). Automated seismic event location by travel-time stacking: An application to mining induced seismicity. *Seismol. Res. Lett., 84*(4), 666-677.

Haddad and Eichhubl, 2020.

Hansen, S. M., & Schmandt, B. (2015). Automated detection and location of microseismicity at Mount St. Helens with a large-N geophone array. *Geophys. Res. Lett., 42(18),* 7390-7397.

Hardebeck, J. L., & Shearer, P. M. (2002). A new method for determining first-motion focal mechanisms. *Bull. Seism. Soc. Am., 92*(6), 2264-2276.

Hardebeck, J. L., & Shearer, P. M. (2003). Using S/P amplitude ratios to constrain the focal mechanisms of small earthquakes. *Bull. Seism. Soc. Am., 93*(6), 2434-2444.

Harris, R. A. (1998). Forecasts of the 1989 Loma Prieta, California, earthquake. *Bulletin of the Seismological Society of America, 88(4),* 898-916.

Hennings, P. H., Lund Snee, J. E., Osmond, J. L., DeShon, H. R., Dommisse, R., Horne, E., ... & Zoback, M. D. (2019). Injection-induced seismicity and fault-slip potential in the Fort Worth Basin, Texas. *Bulletin of the Seismological Society of America, 109(5),* 1615-1634.

Hennings, P. H., Nicot, J. P., Gao, S., DeShon, H. R., Lund Snee, J. E., Morris, A. P., Brudzinski, M. R., Horne, E. A., and C. Breton (2021). Pore pressure threshold and fault slip potential for induced earthquakes in the 6 Dallas-Fort Worth area of north central Texas. *Geophysical Research Letters,* 48, e2021GL093564. https://doi.org/10.1029/2021GL093564.

Holland, A. A. (2013). Earthquakes triggered by hydraulic fracturing in south-central Oklahoma. *Bull. Seism. Soc. Am., 103*(3), 1784-1792.

Hornbach, M. J., DeShon, H. R., Ellsworth, W. L., Stump, B. W., Hayward, C., Frohlich, C., … Luetgert, J. H. (2015). Causal factors for seismicity near Azle, Texas, *Nature Communications, 6*. doi: 10.1038/ncomms7728.

Hornbach, M. J., Jones, M., Scales, M., DeShon, H. R., Magnani, M. B., Frohlich, C., … Layton, M. (2016). Ellenburger wastewater injection and seismicity in North Texas, *Physics of the Earth and Planetary Interiors, 261,* 54-68. doi:10.1016/j.pepi.2016.06.012.

Horne, E. A., Hennings, P. H., Osmond, J. L., & DeShon, H. R. (2020). Structural characterization of potentially seismogenic faults in the Fort Worth Basin. *Interpretation, 8(2)*, T323-T347.

Horton, S. (2012). Disposal of hydrofracking waste fluid by injection into subsurface aquifers triggers earthquake swarm in central Arkansas with potential for damaging earthquake. *Seismological Research Letters, 83*(2), 250-260. Doi:10.1785/gssrl.83.2.250.

King Hubbert, M., & Rubey, W. W. (1959). Role of fluid pressure in mechanics of overthrust faulting: I. Mechanics of fluid-filled porous solids and its application to overthrust faulting. *Geological Society of America Bulletin, 70(2)*, 115-166.

Inc., C. (2020). *COMSOL*. Retrieved from http://www.comsol.com/products/multiphysics/

Ishii, M., Shearer, P. M., Houston, H., & Vidale, J. E. (2005). Extent, duration and speed of the 2004 Sumatra–Andaman earthquake imaged by the Hi-Net array. *Nature, 435*(7044), 933-936.

Janska, E., & Eisnet, L. (2012). Ongoing seismicity in the Dallas-Fort Worth area. *The Leading Edge, 31*912), 1462-1468. Doi:10.1190/tle31121462.1.

Johnson, K. S. (2008). Geologic history of Oklahoma. *Oklahoma Geological Survey Educational Publication, 9(8).*

Justinic, A. H., Stump, B., Hayward, C., and Frohlich, C. (2013). Analysis of the Cleburne, Texas, earthquake sequence from June 2009 to June 2010, *Bulletin of the Seismological Society of America*, *103*(6), 3083-3093. doi:10.1785/0120120336.

Kagan, Y. Y. (2005). Double-couple earthquake focal mechanism: random rotation and display. *Geophys. J. Int.*, *163*(3), 1065-1072.

Keller, G. R., & Hatcher Jr, R. D. (1999). Some comparisons of the structure and evolution of the southern Appalachian-Ouachita orogeny and portions of the Trans-European Suture Zone region. *Tectonophysics, 314*(1-3), 43-68. doi:10.1016/S0040-1951(99)00236-X.

Kennett, B. L. N., & Engdahl, E. R. (1991). Traveltimes for global earthquake location and phase identification. *Geophysical Journal International, 105(2)*, 429-465.

Keranen, K. (2016). *Seismicity near the Nemaha fault in northern Oklahoma*. International Federation of Digital Seismograph Networks. https://doi.org/10.7914/SN/XR_2016.

Keranen, K. M., & Weingarten, M. (2018). Induced seismicity. *Annual Review of Earth and Planetary Sciences, 46*, 149-174. Doi:10.1146/annurev-earth-082517-010054.

Keranen, K. M., Savage, H. M., Abers, G. A., & Cochran, E. S. (2013). Potentially induced earthquakes in Oklahoma, USA: Links between wastewater injection and the 2011 Mw 5.7 earthquake sequence. *Geology, 41(6)*, 699-702.

Keranen, K. M., Weingarten, M., Abers, G. A., Bekins, B. A., & Ge, S. (2014). Sharp increase in central Oklahoma seismicity since 2008 induced by massive wastewater injection. *Science*, *345*(6195), 448-451. Doi:10.1126/science.1255802.

Kiser, E., & Ishii, M. (2017). Back-projection imaging of earthquakes. *Annu. Rev. Earth Planet Sci.*, *45*, 271-299.

Kong, Q., Inbal, A., Allen, R. M., Lv, Q., & Puder, A. (2019). Machine learning aspects of the MyShake global smartphone seismic network. *Seismol. Res. Lett.*, *90*(2A), 546-552.

Koulakov, I. (2009). Out-of-network events can be of great importance for improving results of local earthquake tomography. *Bulletin of the Seismological Society of America, 99(4)*, 2556-2563.

Li, Z., Peng, Z., Hollis, D., Zhu, L., & McClellan, J. (2018a). High-resolution seismic event detection using local similarity for Large-N arrays. *Sci. Rep.*, *8*(1), 1-10.

Li, Z., Meier, M. A., Hauksson, E., Zhan, Z., & Andrews, J. (2018b). Machine learning seismic wave discrimination: Application to earthquake early warning. *Geophys. Res. Lett.*, *45*(10), 4773-4779.

Loucks, R. G., Reed, R. M., Ruppel, S. C., & Jarvie, D. M. (2009). Morphology, genesis, and distribution of nanometer-scale pores in siliceous mudstones of the Mississippian Barnett Shale. *Journal of sedimentary research, 79(12)*, 848-861.

Liu, L., & Zoback, M. D. (1997). Lithospheric strength and intraplate seismicity in the New Madrid seismic zone. *Tectonics, 16(4)*, 585-595.

Lund Snee, J. E., & Zoback, M. D. (2016). State of stress in Texas: Implications for induced seismicity. *Geophysical Research Letters*, *43*(19). Doi:10.1002/2016GL070974.

Magnani, M. B., Blanpied, M. L., DeShon, H. R., and Hornbach, M. J. (2017). Discriminating between natural vs. induced seismicity from long-term deformation history of intraplate faults, *Science Advances, 3(11),* doi:10.1126/sciadv.1701593.

Nuttli, O. W. (1973). Seismic wave attenuation and magnitude relations for eastern North America, *J. Geophys. Res., 78*, 876-885. doi:10.1029/JB078i005p00876.

Ogwari, P. O., DeShon, H. R., and Hornbach, M. J. (2018). The Dallas-Fort-Worth airport earthquake sequence: Seismicity beyond injection period, *Journal of Geophysical Research, Solid Earth 123,* doi:10.1002/2017JB015003.

Paige, C. C., and Saunders, M. A. (1982). LSQR: An algorithm for sparse linear equations and sparse least squares. *ACM transactions on Mathematical Software, 8(1),* 43-71.

Patton, J. M., Guy, M.R. Benz, H.M., Buland, R. P., Erickson, B. K., and Kragness, D. S. (2016). Hydra—The National Earthquake Information Center's 24/7 Seismic Monitoring, Analysis, Catalog Production, Quality Analysis, and Special Studies Tool Suite, *U.S. Geological Survey Open-File Report 2016-1128,* 28 p., doi:10.3133/ofr20161128.

Pavlis, G. L., Vernon, F., Harvey, D., and Quinlan, D. (2004). The generalized earthquake-location (GENLOC) package: An earthquake-location library, *Computers and Geosciences, 30*(9-10), 1079-1091. doi:10.1016/j.cageo.2004.06.010.

Pei, S., Peng, Z., & Chen, X. (2018). Locations of injection-induced earthquakes in Oklahoma controlled by crustal structures. *J. Geophys. Res. Solid Earth, 123*(3), 2332-2344.

Peng, Z., & Zhao, P. (2009). Migration of early aftershocks following the 2004 Parkfield earthquake. *Nature Geoscience, 2(12),* 877-881.

Perol, T., Gharbi, M., & Denolle, M. (2018). Convolutional neural network for earthquake detection and location. *Science Advances, 4*(2), e1700578.

Pollastro, R. M., Jarvie, D. M., Hill, R. J. and Adams, C. W. (2007). Geologic framework of the Mississippian Barnett Shale, Barnett-Paleozoic total petroleum system, Bend Arch-Fort Worth Basin, Texas, *AAPG bulletin, 91*(4), 405-436. doi:10.1306/10300606008.

Pratt, W. E., & Johnson, D. W. (1926). Local subsidence of the Goose Creek oil field. *The Journal of Geology, 34(7, Part 1),* 577-590.

Qin, Y., Chen, X., Walter, J. I., Haffener, J., Trugman, D. T., Carpenter, B. M., ... & Kolawole, F. (2019). Deciphering the stress state of seismogenic faults in Oklahoma and southern Kansas based on an improved stress map. *Journal of Geophysical Research: Solid Earth, 124(12),* 12920-12934.

Quinones, L. A., DeShon, H. R., Magnani, M. B., and Frohlich, C. (2018). Stress Orientations in the Fort Worth Basin, Texas, Determined from Earthquake Focal Mechanisms, *Bulletin of the Seismological Society of America, 108(3A),* 1124-1132. doi:10.1785/0120170337.

Quinones, L., DeShon, H. R., Jeong, S., Ogwari, P., Sufri, O., Holt, M. M., & Kwong, K. B. (2019). Tracking Induced Seismicity in the Fort Worth Basin: A Summary of the 2008–2018 North Texas Earthquake Study Catalog. *Bull. Seism. Soc. Am., 109*(4), 1203-1216.

Reiter, D., Leidig, M., Yoo, S. H., & Mayeda, K. (2012). Source characteristics of seismicity associated with underground wastewater disposal: A case study from the 2008 Dallas-Fort Worth earthquake sequence. *The Leading Edge, 31*(12), 1454-1460. Doi:10.1190/tle31121454.1.

Reyes, C. G., & West, M. E. (2011). The waveform suite: A robust platform for manipulating waveforms in MATLAB. *Seismological Research Letters, 82(1),* 105–110. https://doi.org/10.1785/gssrl.82.1.104.

Richter, C. F. (1935). An instrumental earthquake magnitude scale, *Bulletin of the Seismological Society of America, 25(1),* 1-32.

Ross, Z. E., Meier, M. A., Hauksson, E., & Heaton, T. H. (2018). Generalized seismic phase detection with deep learning. *Bull. Seism. Soc. Am., 108*(5A), 2894-2901.

Ross, Z. E., Trugman, D. T., Hauksson, E., & Shearer, P. M. (2019). Searching for hidden earthquakes in Southern California. *Science, 364*(6442), 767-771.

Rubinstein, J. L., Ellsworth, W. L., McGarr, A., & Benz, H. M. (2014). The 2001–present induced earthquake sequence in the Raton Basin of northern New Mexico and southern Colorado. *Bulletin of the Seismological Society of America, 104*(5), 2162-2181. Doi:10.1785/0120140009.

Rubinstein, J. L., & Mahani, A. B. (2015). Myths and facts on wastewater injection, hydraulic fracturing, enhanced oil recovery, and induced seismicity. *Seismological Research Letters, 86(4),* 1060-1067.

Rubinstein, J. L., Ellsworth, W. L., & Dougherty, S. L. (2018). The 2013–2016 Induced Earthquakes in Harper and Sumner Counties, Southern Kansas. *Bulletin of the Seismological Society of America, 108*(2), 674-689. Doi:10.1785/0120170209.

Savvaidis, A., Young, B., Huang, D.-G., and Lomax, A. (2019). TexNet: A Statewide Seismological Network in Texas, *accepted at Seismological Research Letters, 2019*.

Savvaidis, A., Lomax, A., & Breton, C. (2020). Induced Seismicity in the Delaware Basin, West Texas, is Caused by Hydraulic Fracturing and Wastewater Disposal. *Bull. Seism. Soc. Am., 110*(5), 2225-2241.

Scales, M., DeShon, H. R., Magnani, M. B., Walter, J. I., Quinones, L., Pratt, T. L., and Hornbach, M. J. (2017). A decade of induced slip on the causative fault of the 2015 M_w4.0 Venus earthquake, northeast Johnson County, Texas, *Journal of Geophysical Research: Solid Earth, 122*, 7879-7894, doi: 10.1002/2017JB014460.

Schoenball, M., & Ellsworth, W. L. (2017). A systematic assessment of the spatiotemporal evolution of fault activation through induced seismicity in Oklahoma and southern Kansas. *J. Geophys. Res. Solid Earth, 122*(12), 10-189.

Segall, P. (1989). Earthquakes triggered by fluid extraction. *Geology, 17(10),* 942-946.

Segall, P., Grasso, J. R., & Mossop, A. (1994). Poroelastic stressing and induced seismicity near the Lacq gas field, southwestern France. *Journal of Geophysical Research: Solid Earth, 99(B8)*, 15423-15438.

Segall, P., & Lu, S. (2015). Injection-induced seismicity: Poroelastic and earthquake nucleation effects. *Journal of Geophysical Research: Solid Earth, 120*(7), 5082-5103. Doi:10.1002/2015JB012060.

Shelly, D. R., Beroza, G. C., & Ide, S. (2007). Non-volcanic tremor and low-frequency earthquake swarms. *Nature, 446(7133)*, 305-307.

Skoumal, R. J., Brudzinski, M. R., Currie, B. S., & Levy, J. (2014). Optimizing multi-station earthquake template matching through re-examination of the Youngstown, Ohio, sequence. *Earth Planet Sci. Lett., 405*, 274-280.

Skoumal, R. J., Brudzinski, M. R., & Currie, B. S. (2015). Earthquakes induced by hydraulic fracturing in Poland Township, Ohio. *Bull. Seism. Soc. Am., 105*(1), 189-197.

Skoumal, R. J., Ries, R., Brudzinski, M. R., Barbour, A. J., & Currie, B. S. (2018). Earthquakes induced by hydraulic fracturing are pervasive in Oklahoma. *J. Geophys. Res. Solid Earth*, *123*(12), 10-918.

Skoumal, R. J., Brudzinski, M. R., Currie, B. S., & Ries, R. (2020). Temporal patterns of induced seismicity in Oklahoma revealed from multi-station template matching. *Journal of Seismology, 24*(5), 921-935.

Simpson, D. W., Leith, W. S., & Scholz, C. H. (1988). Two types of reservoir-induced seismicity. *Bulletin of the Seismological Society of America, 78(6)*, 2025-2040.

Smye, K. M., Lemons, C. R., Eastwood, R., McDaid, G., and Hennings, P. H. (2019). Stratigraphic Architecture and Petrophysical Characterization of Formations for Deep Disposal in the Fort Worth Basin, TX. *accepted at Interpretation*, 2019.

Sufri, O., P. Ogwari, L. Quinones, and H.R. DeShon (2018). High Resolution Seismogenic Fault Imaging using Double Difference Relocation in Azle-Reno, TX, 2018 Fall Meeting, American Geophysical Union, Washington DC.

Sweet, J. R., Anderson, K. R., Bilek, S., Brudzinski, M., Chen, X., DeShon, H., ... & Woodward, R. L. (2018). A community experiment to record the full seismic wavefield in Oklahoma. *Seismol. Res. Lett., 89*(5), 1923-1930.

Taira, T. A., Nayak, A., Brenguier, F., & Manga, M. (2018). Monitoring reservoir response to earthquakes and fluid extraction, Salton Sea geothermal field, California. *Science advances, 4(1)*, e1701536.

Talwani, P. (1997). On the nature of reservoir-induced seismicity. *Pure and applied Geophysics, 150(3)*, 473-492.

Waldhauser, F., & Ellsworth, W. L. (2000). A double-difference earthquake location algorithm: Method and application to the northern Hayward fault, California. *Bulletin of the Seismological Society of America, 90(6),* 1353–1368. https://doi.org/10.1785/0120000006.

Walsh III, F. R., & Zoback, M. D. (2016). Probabilistic assessment of potential fault slip related to injection-induced earthquakes: Application to north-central Oklahoma, USA. *Geology, 44(12),* 991-994.

Walter, J. I., Ogwari, P., Thiel, A., Ferrer, F., Woelfel, I., Chang, J. C., ... & Holland, A. A. (2020). The Oklahoma Geological Survey Statewide Seismic Network. *Seismol. Res. Lett.*, *91*(2A), 611-621.

Walter, J. I., Ogwari, P., Thiel, A., Ferrer, F., & Woelfel, I. (2021). EasyQuake: Putting machine learning to work for your regional seismic network or local earthquake study. *Seismological Society of America*, *92*(1), 555-563.

Wang, H. F. (2017). Theory of linear poroelasticity with applications to geomechanics and hydrogeology. Princeton University Press.

Weingarten, M., Ge, S., Godt, J. W., Bekins, B. A., & Rubinstein, J. L. (2015). High-rate injection is associated with the increase in US mid-continent seismicity. *Science*, *348*(6241), 1336-1340. doi:10.1126/science.aab1345.

Wiemer, S., & Wyss, M. (2000). Minimum magnitude of completeness in earthquake catalogs: Examples from Alaska, the western United States, and Japan. *Bulletin of the Seismological Society of America*, *90*(4), 859-869. Doi:10.1785/0119990114.

Woessner, J., & Wiemer, S. (2005). Assessing the quality of earthquake catalogues: Estimating the magnitude of completeness and its uncertainty. *Bulletin of the Seismological Society of America*, *95*(2), 684-698. Doi:10.1785/0120040007.

Yoon, C. E., Huang, Y., Ellsworth, W. L., & Beroza, G. C. (2017). Seismicity during the initial stages of the Guy-Greenbrier, Arkansas, earthquake sequence. *J. Geophys. Res. Solid Earth*, *122*(11), 9253-9274.

Zhai, G., & Shirzaei, M. (2018). Fluid injection and time-dependent seismic hazard in the Barnett Shale, Texas. *Geophysical Research Letters*, *45*(10), 4743-4753. Doi:10.1029/2018GL077696.

Zhai, G., Shirzaei, M., Manga, M., & Chen, X. (2019). Pore-pressure diffusion, enhanced by poroelastic stresses, controls induced seismicity in Oklahoma. *Proc. Natl. Acad. Sci. Unit States Am.*, *116*(33), 16228-16233.

Zhang, H., and Thurber, C.H. (2003). Double-difference tomography: the method and its application to the Hayward Fault, California. *Bulletin of the Seismological Society of America, 93(5)*, 1875-1889.

Zhu, W., & Beroza, G. C. (2019). PhaseNet: a deep-neural-network-based seismic arrival-time picking method. *Geophys. J. Int.*, *216*(1), 261-273.

Zoback, M. D., & Hickman, S. (1982). In situ study of the physical mechanisms controlling induced seismicity at Monticello Reservoir, South Carolina. *Journal of Geophysical Research: Solid Earth, 87(B8)*, 6959-6974.